

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 28, 2002, 12:10:01 ; Search time 3887.7 Seconds
(without alignments)
11758.542 Million cell updates/sec

Title: US-09-674-824-1
Perfect score: 2771
Sequence: 1 cgcacatccactgcctgtg.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824859755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo.hum.*
- 31: em.htgo.inv.*
- 32: em.htgo.rod.*
- 33: em.htg.hum.*
- 34: em.htg.inv.*
- 35: em.htg.rod.*
- 36: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2771	100.0	2771	6	AX010492	AX010492 Sequence
2	2541.8	91.7	2662	6	AX031272	AX031272 Sequence
3	2519	90.9	2591	8	AF091803	AF091803 Triticum
4	2486.6	89.7	2575	8	TAE292521	TAE292521 Triticum
5	2332.8	84.2	2421	8	TAE292522	TAE292522 Triticum
6	2239	80.8	2239	6	A93350	A93350 Sequence 1
7	2054	74.1	2055	8	TAU48227	TAU48227 Triticum ae
8	1393	50.3	2383	8	AF036891	AF036891 Sequence 1
9	1366	49.3	2391	8	AF036891	AF036891 Zea mays
10	1364.4	49.2	2592	8	AF168786	AF168786 Sorghum b
11	1332	48.1	2590	6	AR049919	AR049919 Sequence
12	1329.4	48.0	2533	6	E06904	E06904 cDNA encodi
13	1329.4	48.0	2533	6	RICSSS	RICSSS
14	1329.4	48.0	2533	6	AF165890	AF165890 Oryza sat
15	1327.6	47.9	2592	8	AF63308	AF63308 Sequence 1
16	1325.6	47.8	1752	6	AR106494	AR106494 Sequence
17	1248.4	45.1	1620	6	AR106498	AR106498 Sequence
18	752.8	27.2	2360	6	A51207	A51207 Sequence 9
19	752.8	27.2	2360	6	AR112888	AR112888 Sequence
20	752.8	27.2	2360	6	STSTASYNT	STSTASYNT
21	648.4	23.4	1758	6	A51201	A51201 S. tuberosum
22	648.4	23.4	1758	6	AR112885	AR112885 Sequence 3
23	596	21.5	11387	8	AF091802	AF091802 Aegilops
24	567.6	20.5	10336	6	AX031274	AX031274 Sequence
25	431	15.6	8387	8	AF234163	AF234163 Hordeum v
26	349	12.6	2380	6	AF049921	AF049921 Sequence 1
27	349	12.6	2478	6	A63310	A63310 Sequence 3
28	299	10.8	5072	6	AX031276	AX031276 Sequence 7
29	287	10.4	2793	6	A51205	A51205 Sequence 7
30	287	10.4	2793	6	AR112887	AR112887 Sequence
31	284.4	10.3	2325	8	AF068834	AF068834 Ipomoea b
32	283.2	10.2	2988	8	PSSTASYNT	PSSTASYNT
33	279	10.1	1926	6	A51203	A51203 Sequence 5
34	279	10.1	1926	6	AR112886	AR112886 Sequence
35	275.6	9.9	2903	8	AF383878	AF383878 Oryza sat
36	264.2	9.5	8900	8	RICE1	RICE1
c 37	284.2	9.5	170371	8	AB026295	AB026295 Oryza sat
c 38	255.8	9.2	2595	8	STSTARSYN	STSTARSYN
c 39	253.8	9.2	256191	2	AC084046	AC084046 Trypanoso
40	252.6	9.1	93735	8	ATAC008261	ATAC008261 Arabidops
41	249.4	9.0	2575	8	AF173900	AF173900 Manihot e
42	248.2	9.0	2394	8	AF395537	AF395537 Oryza sat
43	233	8.4	2097	6	AR106493	AR106493 Sequence
44	233	8.4	2480	8	AF019297	AF019297 Zea mays
45	232	8.4	2865	8	AF019296	AF019296 Zea mays

ALIGNMENTS

RESULT 1	AX010492	2771 bp	DNA	PAT	06-SEP-2000
LOCUS	Sequence 1 from Patent WO9958688.				
DEFINITION	AX010492				
ACCESSION	AX010492.1	GI:9997335			
VERSION					
KEYWORDS	bread wheat.				
SOURCE	Triticum aestivum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
REFERENCE	1 (bases 1 to 2771)				
AUTHORS	Luetticke,S., Block,M. and Loerz,H.				
TITLE	Nucleic acid molecules which code for enzymes derived from wheat and which are involved in the synthesis of starch				
JOURNAL	Patent: WO 9958688-A 1 18-NOV-1999;				
FEATURES	LUETTICKE STEPHANIE (DE); BLOCK MARTINA (DE); LOERZ HORST (DE); HOECHST SCHERING AGREVO GMBH (DE) Location/Qualifiers				

[illegible]

QY	1741	atgcttggatctggggatccaatttttgaagctgagatgaatctacagtcagtcagttac	1800
Db	1741	ATGCTTTGGATCTGGGGATCCAATTTTGAAGCTGGATGAGATCTACCGAGTCGAGTTAC	1800
QY	1801	aaggataaattccgtggatgggttggattagtttccagtttccccacagaataactgca	1860
Db	1801	AAGGATAAATTCCTGGGATGGGTGGATTTAGTTGTTCCAGTTTCCACAGNAATACTGCA	1860
QY	1861	ggttcgatattgtttaatgccatcgagatttgaaccttgccggttccatcagctatat	1920
Db	1861	GGTTCGGATATATTGTTAATGCCATCGAGATTGAACTTCGCGTCTTAATCAGCTATAT	1920
QY	1921	gctatgcaatgtatcagttccctgttagtttoatggaactggggcctccgagacacagtc	1980
Db	1921	GCTATGCAATATGTGTACAGTTTCCTGTAGTTTCATGGAACTGGGGCCTCCGAGACAGTCT	1980
QY	1981	gagaccttcaaccttttggctgcaaaaggagagaggtacagggctggcgcttctccacg	2040
Db	1981	GAGACCTTCAACCCCTTTTGGTGCAAAAGAGAGGAGGGGTACAGGGTGGCGTCTCACGG	2040
QY	2041	ctaaccgtgacaagatgttgtggcattgcygaaccgcgatgtcgacattcaggaggac	2100
Db	2041	CTAACCCTGCACAAGATGTTGTGGCATTTGCGAACCCTGATGTGACATTCAGGAGCAC	2100
QY	2101	aagccgtccctgggaggggctcatgaagcgaggcatgacaaagaacatacgtggagccat	2160
Db	2101	AAGCCGTCTTGGAGGGGCTCATGAAGCGAGGCATGACAAAGACCATACGTGGGACCAT	2160
QY	2161	gccccagacagtcacgacagatcttcgaagcgaggtgggcttcgtgaccaaaccctacgtc	2220
Db	2161	GCCCCGAGCAGTACGAGCAGATCTTCGAGTGGGCCCTTCGTGACCACCCCTACGTCATGT	2220
QY	2221	agacggggactggggaggttccaaagtcgaggtctccttcagctcttgaagaatcctttca	2280
Db	2221	AGACGGGACTGGGAGGTCCAAGTCCGAGTCTCCTTCAGTCTGGAACATCCTCTTCA	2280
QY	2281	tccttcgcgcgcggcgaagataccctctgacattgcgttcctgctacagtagagtcg	2340
Db	2281	TCCTTCGCGCGCCGGAAGGATACCCCTGTACATTTGCGTTCTCTGTGTACAGTAGAGTCG	2340
QY	2341	caatgcgcctgcttgcgttgcgcgggttcgagaacatacgcgctgctgctgcgcg	2400
Db	2341	CAATGCGCCTGTCTGCTTTGGTTTGCCTGGTTCGCCGGTTCGAGAACATATGACGGCTGTGCTGCTGCG	2400
QY	2401	gcggtgcagcgttcgggtggacgacagttacagttttggggataaaggaaaggatgct	2460
Db	2401	CGGCTGACAGCTTCGGGTGGACGACAGTTACAGTTTGGGGAATAAGGAAGGATGTGCT	2460
QY	2461	gcagatggtttaacagcaaaagcaccatcgatggcgacctctctgcgtgttaccagct	2520
Db	2461	GCAGATGGTTTAAACAGCAAAAGCACCATCAGATGGCAGCCTCTCTGTCGCTGTACAGCT	2520
QY	2521	gaatcagaaccactggtagctctttagccttagtgattggaagtttgtgcctct	2580
Db	2521	GAAATCAGAAACCAACTGGTGACTCTTTACCTTTAGTGATTGTGAAGTTGTGCTTCT	2580
QY	2581	gtgcatgtgcttgccttagctgacaaatatttgacctgttggagaattttatcttcg	2640
Db	2581	GTGTATGTTGTTGCTTCTTAGCTGACAAATATTTCACCTGTTGGAGAATTTATCTTTG	2640
QY	2641	ctgcgttttttttaacaaagaggggtttccctccgatttcatttaaaaaa	2700
Db	2641	CTGCTGTTTTTTTTTAAATCAAAAGAGGGGTTTCTCCGATTTCATTAATAAAAAA	2700
QY	2701	aa	2760
Db	2701	AA	2760
QY	2761	aaaaaaaaaaaa 2771	
Db	2761	AAAAAAAAAAAA 2771	

RESULT	2
AX031272	
LOCUS	AX031272 2662 bp DNA PAT 20-SEP-2000
DEFINITION	Sequence 11 from Patent WO9914314.
ACCESSION	AX031272
VERSION	AX031272.1 GI:10278604
KEYWORDS	Aegilops tauschii.
SOURCE	Aegilops tauschii.
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Aegilops.
REFERENCE	1 (bases 1 to 2662)
AUTHORS	Li,Z., Morell,M. and Rahman,S.
TITLE	Regulation of gene expression in plants
JOURNAL	Patent: WO 9914314-A 11 25-MAR-1999;
	GOODMAN FELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;
	RAHMAN SADEQUER (AU) ; UNIV AUSTRALIAN (AU) ; COMMW SCIENT IND RES
FEATURES	ORG (AU) : GROUPE LIMAGRAIN PACIFIC PTY L (AU) Location/Qualifiers 1..2662 /organism="Aegilops tauschii" /db_xref="taxon:37682" /tissue_type="ENDOSPERM"
misc_feature	1..2651
BASE COUNT	592 a 677 c 754 g 639 t
ORIGIN	/product="NUCLEOTIDE SEQUENCE OF CDNA WHEAT SSS I"
Query Match	91.7%; Score 2541.8; DB 6; Length 2662;
Best Local Similarity	98.6%; Pred. No. 0;
Matches 2597;	Conservative 0; Mismatches 32; Indels 6; Gaps 3;
Qy	25 ctccacttctctcccgcgcacaccgagtgcggcaccgtctcatcacccatcacctcg 84
Db	
2	CTCCACTCTTCTCTCCCGCGCACACCAGTCGGCACC GGCTCATCACTACCTGCG 61
Qy	85 gcttcggccaccggaacccccccgatcgttttgaggcagcgacctaaacccccggg 144
Db	62 GCCTCGGCCACGGCAAACCCCCGCATCGCTTTTGACGAGCAGCGCACTAAACCCCGGG 121
Qy	145 gagcgccccggcgagcagcagcagcagcagtcggtggagagagaggtcttcgcccgcccg 204
Db	122 GACGCGCCCCGCGGCAGCAGCAGCAGCAGCAGTCGGAGAGAGAGGCTTCGCCCGGCCCG 181
Qy	205 caccgagcgggcgatccacgctcgttcgctccacactctcgcctctctcccttgtcc 264
Db	182 CACCGAGCGGGCGATCCACCGTCGGTCGTCCCGACCTCTCTCGGCCTCTCTCCCTGTCC 241
Qy	265 cgcgcccccaaccatcggcgacggcggtgcggcgccgggtgcctcgccccagcgctc 374
Db	242 CGCGCGCCCACACCCATGGCGGCGACGGGCGTCGCGCGCGGTGCCTTCGCCCGCCACGTC 301
Qy	325 cgcctgcgcgcgatccggcgacgcggcgccggcgagctcgcctcgttcgcgcgcgg 384
Db	302 CGCCTGCGCGCGATCCGGCGACGGCGCGCGCGGCGTCCGCCCTGCTGTCGCGCGCGG 361
Qy	385 ctccgcgcttggcgcgggcgccgtacgtgcgcgagctcagcaggagggcgcccgcgcg 444
Db	362 CTCGGCGCTTTGGCGCGGGGCGCGCTAGCTTTCGCGAGCTACGAGGAGGCGCCCGCGCG 421
Qy	445 gcggcgcgagcagcagaactggccccgcgctcgtgccagggttcctcgcgcgcg 504
Db	422 CGCCCCGCGAGCAGCAGCAACTGGCCCCCGCCGCTCGTCCAGGCTTCCTCGCGCGCGG 481
Qy	505 ccgccccgcgcgcgaagtccgcggccccgcgacgcagccgcccccttcgcgcgcgcg 564
Db	482 CCGCCCGCGCGCCCGCAGTCGCGCGCGCCCGCAGCAGCGCGCCCCCTCCCGGAGCGCGCGGTG 541
Qy	565 gggaactcgcgcgcgacctctcgtcgaaggattgtaggattccatcagcagcata 624
Db	542 GGGGAATCGCGCCGCACTCTCTCTGAAGGGATTGCTGAGGATTTCATCGACAGCATA 601

QY	625	atttggtgctcaagtagagcaggtgcaatttgcattgctgagatcattgagcgaaccccaagct	684	1705	attccagagctcatagagagcagctgcaatttgcattgctgagatcattgagcgaaccccaatt	1764
Db	602	atttggtgctcaagtagagcaggtgcaatttgcattgctgagatcattgagcgaaccccaagct	661	1682	attccagagctcatagagagcagctgcaatttgcattgctgagatcattgagcgaaccccaatt	1741
QY	685	aaagttcacagtagcagtcggttggtagcaggtggaagcagcctcctcattgcaaaagtcagg	744	1765	tttgaaggtcgatgagatctccagagtcagattccagagataaaattccgtagagtggtt	1824
Db	662	aaagttcacagtagcagtcggttggtagcaggtggaagcagcctcctcattgcaaaagtcagg	721	1742	tttgaaggtcgatgagatctccagagtcagattccagagataaaattccgtagagtggtt	1801
QY	745	gggttgagagatgttgggttgcttaccgaatgctcttgcctgctgctggtcaccagtg	804	1825	ggatttagttccagtttcccaagataaactgcaggtgtgcgatatattgtttaatgcc	1884
Db	722	gggttgagagatgttgggttgcttaccgaatgctcttgcctgctgctggtcaccagtg	781	1802	ggatttagttccagtttcccaagataaactgcaggtgtgcgatatattgtttaatgcc	1861
QY	805	atggttgtaatgcgaagatactaaatgggtccctcctgataaaaactatgcaaaagcatta	864	1885	tcgagattgaaacctgctgctttaaactagctatactgatacaataatggtacagctcct	1944
Db	782	atggttgtaatgcgaagatactaaatgggtccctcctgataaaaactatgcaaaagcatta	841	1862	tcgagattgaaacctgctgctttaaactagctatactgatacaataatggtacagctcct	1921
QY	865	tacactgcgaagcacatlaagattccatgcttcttgggggatacacatgaagtgaactttttt	924	1945	gtagttcaatggaactgggggctcccaagacacacagtcgagaccttcaaccttttggtag	2004
Db	842	tacactgcgaagcacatlaagattccatgcttcttgggggatacacatgaagtgaactttttt	901	1922	gtagttcaatggaactgggggctcccaagacacacagtcgagaccttcaaccttttggtag	1981
QY	925	catgagtagagacaacgcgtcgatgggtgttgcagatccgctcattatcacagacca	984	2005	aaagagagagaggttacaggtgtggcgcttctcaccgctaaccgttggaagaatgtgtgg	2064
Db	902	catgagtagagacaacgcgtcgatgggtgttgcagatccgctcattatcacagacca	961	1982	aaagagagagaggttacaggtgtggcgcttctcaccgctaaccgttggaagaatgtgtgg	2041
QY	985	ggaagtttatggagataatttgggtgcttcttgggtgataaactagttccagatacacactc	1044	2065	gcattgcgaacccgcgagtgcgacaattcagggagcacagcgctcctgggggggctcatg	2124
Db	962	ggaagtttatggagataatttgggtgcttcttgggtgataaactagttccagatacacactc	1021	2042	gcattgcgaacccgcgagtgcgacaattcagggagcacagcgctcctgggggggctcatg	2101
QY	1045	cttgcctatgctgcagtcgagggcccaactaaactccttgaattggaggatatttatgga	1104	2125	aaagagagtagcaaaagacacacacagtcaggtgggagccatgoc-ccgagcagtcagcagcagc	2183
Db	1022	cttgcctatgctgcagtcgagggcccaactaaactccttgaattggaggatatttatgga	1081	2102	aaagagagtagcaaaagacacacacagtcaggtgggagccatgoc-ccgagcagtcagcagcagc	2161
QY	1105	cagaattgcaatgttgtgtgaacgattggcagtcgagccttgtgacagtcctcttgcct	1164	2184	ttcgagtggtggtccttcagctcctgaagacacatcctcttcacccgcccgggaagagata	2243
Db	1082	cagaattgcaatgttgtgtgaacgattggcagtcgagccttgtgacagtcctcttgcct	1141	2162	ttcgagtggtggtccttcagctcctgaagacacatcctcttcacccgcccgggaagagata	2221
QY	1165	gcaaaatagacacacacagtggtttacagagatcccgagcagcccttgccttgcataat	1224	2244	gtcgagtcctccttcagctcctgaagacacatcctcttcacccgcccgggaagagata	2303
Db	1142	gcaaaatagacacacacagtggtttacagagatcccgagcagcccttgccttgcataat	1201	2222	gtcgagtcctccttcagctcctgaagacacatcctcttcacccgcccgggaagagata	2281
QY	1225	ttagcacatcagagtgtagcctgcgaagtacatacctgatactggtgagtcgctcctgaa	1284	2304	ccccctacatgctgctcctgctacagtagagtcgaatgcgctcctgcttcttgcct	2363
Db	1202	ttagcacatcagagtgtagcctgcgaagtacatacctgatactggtgagtcgctcctgaa	1261	2282	ccccctacatgctgctcctgctacagtagagtcgaatgcgctcctgcttcttgcct	2340
QY	1285	tgtatgagccttagaattgggtatttccagaatgggcaaggagcagcctcctgacaag	1344	2364	cgccggttcgagaacacatagcggtgctgctgctgcggcggtgcagagcttcgggtggagc	2423
Db	1262	tgtatgagccttagaattgggtatttccagaatgggcaaggagcagcctcctgacaag	1321	2341	cgccggttcgagaacacatagcggtgctgctgctgcggcggtgcagagcttcgggtggagc	2400
QY	1345	ggtgagcaggttaacttttgaagagcaggtgtgtgacagcagatcggattgtgaccgctc	1404	2424	acagttacagtttttgggggaataaaggaggaatgctgcagagatggttaacagcaagca	2483
Db	1322	ggtgagcaggttaacttttgaagagcaggtgtgtgacagcagatcggattgtgaccgctc	1381	2401	acagttacagtttttgggggaataaaggaggaatgctgcagagatggttaacagcaagca	2460
QY	1405	agtcaggggtatttcaatggagagtcacaaactgctgaagtggtgacagcagatcggatgc	1464	2484	ccactcagatgagcagcctcctgctgcggtgttacagctgaaatcagaaacacacgctgac	2543
Db	1382	agtcaggggtatttcaatggagagtcacaaactgctgaagtggtgacagcagatcggatgc	1441	2461	ccactcagatgagcagcctcctgctgcggtgttacagctgaaatcagaaacacacgctgac	2520
QY	1465	ttaagctcccgaaaaagtgattgaaatggaaattgaaatggaaattgacattaatgattgg	1524	2544	tccttagccttagtgattgtaagatttgcgcttcctgctgctgctgctgctgctgctgctg	2603
Db	1442	ttaagctcccgaaaaagtgattgaaatggaaattgaaatggaaattgacattaatgattgg	1501	2521	tccttagccttagtgattgtaagatttgcgcttcctgctgctgctgctgctgctgctgctg	2580
QY	1525	aacccccacacagacaagtgctccctcctcctcctcctcctcctcctcctcctcctcctc	1584	2604	tgacaataatttgacctgttggagaat-----tttatctcttgcgctgctgctgctgctgctg	2654
Db	1502	aacccccacacagacaagtgctccctcctcctcctcctcctcctcctcctcctcctcctc	1561	2581	tgacaataatttgacctgttggagaat-----tttatctcttgcgctgctgctgctgctgctg	2635
QY	1585	gcaaaatgaagcgtgaattgcagagaggttgggtttaccctgtgaaggagagattgcttcc	1644	RESULT	3	
Db	1562	gcaaaatgaagcgtgaattgcagagaggttgggtttaccctgtgaaggagagattgcttcc	1621	AF091803	LOCUS	AF091803 2591 bp mRNA 14-SEP-1999
QY	1645	ctgattggccttttatggaagactggattaccagaaaagcattgactcctattaaatggcc	1704	DEFINITION		Triticum aestivum starch synthase I mRNA, complete cds.
Db	1622	ctgattgg				

ORGANISM *Triticum aestivum*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 2591)
 AUTHORS Li, Z., Rahman, S., Kosar-Hashemi, B., Mouille, G., Appels, R. and
 Morell, M.K.
 TITLE Cloning and characterization of a gene encoding wheat starch
 synthase I
 JOURNAL Theor. Appl. Genet. 98, 1208-1216 (1999)
 REFERENCE 2 (bases 1 to 2591)
 AUTHORS Li, Z., Rahman, S., Kosar-Hashemi, B., Mouille, G., Appels, R. and
 Morell, M.K.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-1998) CSIRO Division of Plant Industry, GPO Box
 1600, Canberra, ACT 2601, Australia
 FEATURES Location/Qualifiers
 source 1..2591
 /organism="Triticum aestivum"
 /db_xref="taxon:4565"
 /chromosome="7"
 /tissue_type="endosperm"
 256..2199
 /codon_start=1
 /product="starch synthase I"
 /protein_id="AAD54661.1"
 /db_xref="GI:5880466"
 /translation="MAATGVGAGCLAPSVRLRADPATAARASACVVRRLRLRLRGRI
 VAESREGPARAAQOOLAPPLVPGLAPPAPAPQAPTOPPLPDAGVGLAFDL
 LLEGIAEDSIDSIIVAAEQDSEIDMNEOQAKVTRISVIVFGEAAPAKSGGLGD
 CGSLPIAALAGHRVVMVPRYLNGSDKNVAKALYTHGHIKIPCRGSHVEYTFHEY
 RNDYVDFVDPHSYRPLSLYGNFAGAGNQRHYLLCTAACAEPLILELGYIYGO
 NCMFVNDHSLVPLVLAARYKPYGVYRDSRLVILNLAHQLEPASTYPLDLGPP
 EYGLALSWPEPARHAKDGEARVNFGLVAVTADRIYVSGYSWEVTTAEGGGL
 NELLSSKSVLNGVINGIDINDWNPDTDKLPHYVSDVLSKAKAKAELOKELGLPV
 REDVPLGLFGRIDYQKIDILKMAIPELMREDYQVMLGSDGPIFEQWWRSTESYK
 DKFNGWGFVSPVSHRITAGCDILLMPRSFPCGLNOLYAMGYVTVPHVGTGGURDT
 VETNPFNGAGEGTGWAFSPITVDKMLWALRTANSTFREHPSPSEGLMKRMKTDHT
 WDHAAEQYEIFEAFVDPQVNM"
 BASE COUNT 567 a 672 c 743 g 609 t
 ORIGIN

Query Match 90.9%; Score 2519; DB 8; Length 2591;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 2559; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

QY 25 ctccactcttctccccgcgcacaccagtcgagtcgacccgctctatccaccatcacctcg 84
 DB 1 CTCCCACTCTCTCTCCCGCGCACACCAGTCCGCCGGCTCATACCCATCACCTCG 60

QY 85 gctcggccacccggaaacccccgcgcacccgttttgcgagcgagcgcaactaaaccccg 144
 DB 61 GCCTCGCCACCGGCAACCCCGCATCGCTTTTCAGGCAGCGCACTAAACCCCGG 120

QY 145 gacgccccgcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 204
 DB 121 GAGCGCGCGCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 180

QY 205 caccgagcgggcgatccaccgttcgctcgccacccctccctccctccctccctccct 264
 DB 181 CACCGAGCGGGGGGATCCACCTCGCTCGCTCGCCACCTCCTCCGCTCCCTCCCT 240

QY 265 cgcgcgccacacccatgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 324
 DB 241 CGCGCGCCACACCATGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 300

QY 325 cgccttcgcgcgcgaatccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 384
 DB 301 CGCCTCGCGCGCGATCCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360

QY 385 ctccgcgcgttcggcgcgggcgagtcagtcgagtcgagtcgagtcgagtcgagtcgagtc 444
 DB 385 ctccgcgcgttcggcgcgggcgagtcagtcgagtcgagtcgagtcgagtcgagtcgagtc 444

DB 361 CTCGCGCGCTTTGGCGGGGGCGCTACGTTGCCGAGCTCAGAGGGAGGCCCGCGGCG 420
 QY 445 cgccccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 504
 DB 421 CGCCCCGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 480
 QY 505 cgccccgc 564
 DB 481 CGCCCCGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540
 QY 565 ggggaactcgc 624
 DB 541 GGGGAACCTCG 600
 QY 625 atgtggtgctgaagtcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 684
 DB 601 ATTGTGGCTGCAAGTAGCAGGAGTCTGAGATCATGATGCGAATGAGCAACTCAAGCT 660
 QY 685 aaagtacacgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 744
 DB 661 AAAGTTACAGCTAGCATCGTGTGTGACTGTGAGGTGAGGTGCTTCTTATGCAAGTCAGG 720
 QY 745 ggttggtgagatgttctggtggttcttaccattgctctgctgctgctgctgctgctgctgctgct 804
 DB 721 GGCTGGGAGATGTTTGTGGTTCGTACCAATTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 805 atggttgaatgccaaagatacttaaatggtcctctgataaaaactatgcaaaagcatta 864
 DB 781 ATGGTTGTAATGCCAAGATACCTTGAATGGGCTCTGATAAAAACTATGCAAGGCATTA 840
 QY 865 tacactggaagacacataagattccattgctcttggggatcacatgaagtgaagtcacattttt 924
 DB 841 TACATGGGAGACACATTAAGATTCCATGCTTGGGGATCACATGAAGTCACCTTTTTT 900
 QY 925 catgagtagagacacacgctgagtggtgttttcgacatcgcgtcatcacacacacacac 984
 DB 901 CATGAGTATAGAGACACGTCGATGGGTGTTGCTCATCATCGTCATATATAGACCA 960
 QY 985 ggaagttatatgagataatttgggtgcttttgggtgataataatcagttcagatcacacac 1044
 DB 961 GAAAGTTTATATGGAGATAATTTTGGTGTCTTTTGGTATAATCAGTTCAGTACACACTC 1020
 QY 1045 cttgtcatgctgcatcgagggcccaactaactctgaattgagggagatatattatgga 1104
 DB 1021 CTTTGTCTATGCTGATCGATCGAGGGCCCCACTATCTTGAATGGGAGGATATATTTATGA 1080
 QY 1105 cagaatgcatgtttgtgtgaacgattggcagcagcctgtgcccagtcctctctgct 1164
 DB 1081 CAGAAATGCGATGTTTGTGTGAACGATTGGCATGCCAGCTTGTGCCACTCTCTTGTCT 1140
 QY 1165 gcaaaatatagacacacacggttttacagagattcccgagcagcaccctgtttatataat 1224
 DB 1141 GCAAAATATAGACCATACGCTGTTTACAGAGATTCGCCGAGCACCCTTGTATTATATAT 1200
 QY 1225 ttacacacacaggggtgtgagcctgcaagtacatactctgactgtggtggtggtggtggtggtggt 1284
 DB 1201 TTAGCACATCAGGGCTGAGGCTGCAAGTACATATCTCTGATCTGGATTGCCACTGAA 1260
 QY 1285 tggatggagctttagaatgggtatttccagaatgggcaagggagcagcagcagcagcagcagcagc 1344
 DB 1261 TGGTATGGAGCTTGAATGGGTATTTCCAGAAATGGGCAAGGAGCATGCCCTTGTGACAAG 1320
 QY 1345 ggtgagcaggttaactttttgaaaggagcagttgtgacagcagcagcagcagcagcagcagcagc 1404
 DB 1321 GGTGAGGCAGTTTAACCTTTTGAAGGAGCAGCTGCTGACAGCAGATCGAATTTGTGACCGTC 1380
 QY 1405 agtcaggggttatctatggaggtcaactgctgaagtgagcagcagcagcagcagcagcagcagcagc 1464
 DB 1381 AGTCAGGGTATTATGAGGAGTCAACTGCTGAAGGTGGACAGGCGCTCAATGAGCTC 1440
 QY 1465 ttaagctccccaaaaagtgtattgaatggaattgtaattgaaattgacattaatgattg 1524
 DB 1441 TTAAGTCCCGAAAAAGTATTGAATGGAATGTAATGGAATGGAATGGAATGGAATGGAATGGAAT 1500

Qy 2604 tgacaaatatt 2614
|||||
Dh 2580 TGACAAATATT 2590

RESULT 4

LOCUS	2575 bp	PLN	21-JUL-2000
DEFINITION	Trifolium aestivum mRNA for starch synthase I-1 (wssi-1 gene).		
ACCESSION	AEJ292521		
VERSION	AJ292521.1		
KEYWORDS	starch synthase I-1; wssi-1 gene.		
SOURCE	bread wheat.		

SOURCE
 ORGANISM
 Triticum aestivum
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Poaceae: Triticeae: Triticum.
 1 (bases 1 to 2575)
 Peng, M., Hucl, P. and Chibbar, R. N.
 Isolation, characterization and expression analysis of starch synthase 1 from wheat (Triticum aestivum L.)
 2 (bases 1 to 2575)
 JOURNAL
 REFERENCE
 2 (bases 1 to 2575)

FEATURES
source

Query Match	89.7%	Score 2486.6;	DB 8;	Length 2575;
Best Local Similarity	98.8%;	pred. No. 0;		
Matches 2526; Conservative	0;	Mismatches	29;	Indels 2; Gaps 2;

Db 718 GCGAAGCACATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTTTCATGAG 777
Qy 931 tatagacaacgctcgattgggtgtttgttcgatcatccctcatatcacagacaggaagt 990
Db 778 TATAGACACAACGTCGATTGGGTGTGTTCGATCATCCCGTCATATACAGACAGGAAGT 837
Qy 991 ttatatggagataattttgtgtctttgtgtgataacacagttcagatcacacactctttgc 1050
Db 938 TTATATGGAGATAAATTTTGGTGTCTTTTGGTGATAATCAGTTACAGATACACACTCTCTTGC 897
Qy 1051 tatgtcatcgcagagcccaactaaaccttgattggaggagatatattttgacagaaat 1110
Db 898 TATGCTGCATGCGAGGCCCCCACTTAATCTTGAAATTTGGGAGGATATATTTATGACAGAAT 957
Qy 1111 tgcattgtttgttgaaacgattggcatgccagccttctgacagtccttcttgcacagaa 1170
Db 958 TGCATGTTTGTGTAACGATTGGCATGCCAGCTTGTGCCAGTCTCTTCTTGTGCAAAA 1017
Qy 1171 tatagaccatagcgttttacagagagattcccgagacaccttggttatataataattagca 1230
Db 1018 TATAGACCATACCGGTTTTACAGAGATTCCCGCAGCACCCCTTGTATACATAAATTAGCA 1077
Qy 1231 catcaggggtgtgagccttcaagatgggcaagagagcattgacacagatcgattgtgacgtcagtcag 1410
Db 1078 CATCAGGGGTGAGACCTGCAAGTACATATCTGTATCTGGGATTGCCCTCTGAAATGGTAT 1137
Qy 1291 gtagctttagaatgggtatttccagaattgggcaagagagcattgacacagatcgattgtgacgtcagtcag 1410
Db 1138 GGAGCTTTAGAAATGGGTATTTCCAGAATGGCAAGGAGCATGCCCTTGACNAGGGTGAG 1197
Qy 1351 gcagtttaacttttgaagagagcagttgtgacacagatcgattgtgacgtcagtcag 1410
Db 1198 GCAGTTAACTTTTGAAGAGGAGCAGTTGTGACAGCAGATCGGATTGTGACCGTCAGTCAG 1257
Qy 1411 ggtattcatgggggtcacactgctgaaggtgagagagcctcaatgacgtcttaagc 1470
Db 1258 GGTATTTCATGGAGGTGCAACATGCTGAAGGTGGACAGGGCCCTCAATGAGCTCTTAAGC 1317
Qy 1471 tcccgaagaaagtattgaatggaaattgtaaatggaattgacattaatgattggaacccc 1530
Db 1318 TCCCGAAAAGTGATTGAATGGAAATTTGTAATGGAATTGACATTAATGATTGGAACCCC 1377
Qy 1531 accacagacaagtctccctcatcattatctgtcgatgacacctctggaagggccaaa 1590
Db 1378 ACCACAGACAAGTCTCCCTCATCATATTCTGTGATGACCTCTCTGGAAGGCCAAA 1437
Qy 1591 tgaagctgaattgcagaaggaggttggtttacctgtaaggagagatttccctctgatt 1650
Db 1438 TGTAAAGCTGAATTGCAAGAGGAGTTGGGTTTACCTGTAAGGAGGATGTTCTCTGATT 1497
Qy 1651 ggtcttattggaagactggattaccagaagcattgatctcaataaaatggccattcca 1710
Db 1498 GGCTTTATTGGAAGACTGGATTACCAAGAAGCATTTGATCTATTAATGATGCCATTCCA 1557
Qy 1711 gacctaagaggagagcgtgcaatttgatcgtctggatcctgggagatcccaattttgaa 1770
Db 1558 GAGCTATGAGGAGGAGACGTGCAATTTGATGCTTGGATCTGGGATCCCAATTTTGA 1617
Qy 1771 ggtggatgagatctaccagctcagttacaagagataaattccgtggatgggttgattt 1830
Db 1618 GGCTGGATGAGATCTACCGAGTCGAGTTACAAGGATAAATCCGTGGATGGGTTGGATT 1677
Qy 1831 agtctccagtttcccacagaataactgcaggttcgatatattgttaattgacacagaga 1890
Db 1678 AGTGTTCAGATTTCACACAGAAATAACTGCAGTTGCGGATATATTGTTAATGCCATCGAGA 1737
Qy 1891 tttagaaccttcggtcttaactcagctatgctatgcgaatggttacagttcctgtagtt 1950
Db 1738 TTTGAACCTTGGGCTTTAATCAGCTATATGCTATGCAATATGTTACAGTTCTCTGAGTT 1797
Qy 1951 catggaactgggggctccgagacacagtcgagaccttcaaccccttttgggtgcaaaagga 2010
|||||

Db 1798 CATGGAACCTGGGGCCTCCGAGACACAGTCGAGACCTTTCAACCTTTTGTGTGCAAAAGGA 1857
Qy 2011 gagagaggtacagaggtggggtgttcacccgtacacgttgacaagatgttgggtgcatg 2070
Db 1858 CAGAGGGGTACAGGGTGGGGTTCACCGCTAACCGTGCACAGATGTTGTGGCATTG 1917
Qy 2071 cgaaccgcgattgcgacattcaggagagacacacccgtccctgggaggggtcctcatgaagca 2130
Db 1918 CGAACCCTGATGTCGACATTCAGGGAGCACACCGCTCTGGGAGGGCTCATGAAGCGA 1977
Qy 2131 ggcatacgaaagaccatacgtgggaccatgccc-cgagcagtcacagcagatcttcag 2189
Db 1978 GGCATGACGAAGAACCATACGTGGGACCATGCGCGCCGAGCAGTACGAGCAGATCTCGAG 2037
Qy 2190 tgggctctcgtgacacacccctcgtcatgacaggggactgggaggtccaaagtgcga 2249
Db 2038 TGGGCTTCTGTCGACCAACCTTACGTATGTAGACGGGACTGGGGAGGTCCAAGTGCA 2097
Qy 2250 gttcctcagctcctgaagacatcctcttcacctcccgccgcccgaagagataccctg 2309
Db 2098 GTCTCCTTGAGCTCTGAAGACATCTCTTTCATCTTCGCGCGCCGGAAGGATACCCCTG 2157
Qy 2310 tacattgcgttgcctgcctacagtagagtcgcaatgcgctgctgtcttgggtgcgag 2369
Db 2158 TACATTTGCGTGTGCTTGTCTTACAGTAGAGTGCATATCGCGCTGCTTGTGCTTGGTTCGCGG 2217
Qy 2370 ttcgaaacatatgacgctgtgctgctgcgagggatgtgctgcaggtggttaacagcaaacccactc 2489
Db 2218 TTTCGAAACATATGACGGCTGTGCTGTGCTGCGCGGTGACAGCTTCGGGTGGACACAGTT 2277
Qy 2430 acagtttggggaataaaggaggatgtgctgcaggtggttaacagcaaacccactc 2489
Db 2278 ACAGTTTGGGGAATAAGGAAGGATGTGCTGCAGAGTGTTAACAGCAAGACACACTC 2337
Qy 2490 agatggcagctcctcgtcctgctgttaccagctgaaatcagaacaaactggtgactctta 2549
Db 2338 AGATGGACCTCTCTGTCGCTGTTACAGCTGAAATCAGAAACCACTGGTACTCTTTA 2397
Qy 2550 gctttagtgattggaagtgtt 2573
Db 2398 GCCTTAGTGATTGTGAAGTTGTTT 2421

RESULT 6

LOCUS A93350 2239 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9745545.
ACCESSION A93350
VERSION A93350.1 GI:6741617
KEYWORDS bread wheat.
SOURCE Triticum aestivum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 2239)
AUTHORS BLOCK, M. and Loefer, H.
TITLE NUCLEIC ACID MOLECULES ENCODING ENZYMES FROM WHEAT WHICH ARE INVOLVED IN STARCH SYNTHESIS
JOURNAL Patent: WO 9745545-A 1 04-DEC-1997;
FEATURES
HOECHST SCHERING AGREVO GMBH (DE); BLOCK MARTINA (DE)
source
1. .2239
Location/Qualifiers
/organism="Triticum aestivum"
/strain="CV. FLORIDA"
/db_xref="taxon:4565"
/clone="TASSS"
/maplotype="CA. 21 D CARYOPSES"
/clone_lib="CDNA LIBRARY IN PBLUESCRIPT SK (-)"
<3. .2018
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB69543.1"

CDS

/db_xref="gi:6741618"
/translation="TOPPLDPAAGVGEIAPDLLEGIABEDSIDSIIVAAASQDSEIMDA
NEPOAKYTRSTVFVTEAAPYAKSGGLDVCGLPDLAALAGHRVVMVPRYLNGSS
DKNYAKLYTRAKHIKPCFGGSHVETPFHEYRDNDVWFVDHPSYHRPGSLYDNFGA
PDGNFRYLCLYACEAPLILELGGYIYGONCMFVNDWHASLVPVLLAAKAPRYGV
YRDSRSTLIHNLHQGVPEASTPDLGLPBNYCALEWVPEWARHALLDKGAVNF
LKGVATADRIYTVVSGVSWETTAEGGQGLNELLSRKSVLNGIVGIDIDNDWNP
DKCLPHHSVDDLSKAKCAELQKELGLPVREDYPLIGFIPVSHRITGACDILKMP
ELMRDVOFMLGSDPIFEQWMRSTESSYKDFRGWGFVSPVSHRITGACDILKMP
SRFPCNLNOLYAMQYTPVHVHGTGLRDYVETFPNFGAKGEGTGWAFSPLVDKM
LWALPTMSTFHEKPSWGLMKRMGTLDHWDHAPSTSRSSGSPSWNPTSCRRGL
GSKCESPALYTSSSSPGPEGYPTLRCPATVESCACILLWFAGSRTYDGCACAAV
TAGGQHLQFWMIRKGCACAGMLTAKHSDGSLSVRVTAEIRNLVTL"

BASE COUNT		611 a	448 c	590 g	590 t
ORIGIN					
Query Match		80.8%; Score 2239; DB 6; Length 2239;			
Best Local Similarity		100.0%; Pred. No. 5e-292;			
Matches 2239; Conservative		0; Mismatches	0; Indels	0; Gaps	0;
QY	533	cgacgagccgcccctgcgagccgctggtggggaactcgcgcgcacccctcctcctcg	592		
DB	1	CGAGCAGCCGCCCTCGCGGACCGCGCTGGGGAACTCGCGCCGACCTCCTGCTCG	60		
QY	593	aagggattcgtgaggtatccatcacagcataattggtcgtcaagtgcagcaggtatctg	652		
DB	61	AAGGGATTGCTGAGGATTCCATCAGCAGCATAAATGTGGCTGCAAGTGACAGGATTCTG	120		
QY	653	agatcatgagtcgcaatgagcaactcaagctaaagtacacatgacatcgtgttttgtga	712		
DB	121	AGATCATGGATCGGAATCAGCAACTCAAGCTAAAGTTACACGTAGCATCGTGTGTGA	180		
QY	713	ctggtgaagctcgtcctctatgcaaatgcagggggtggagagattgttggttcgttac	772		
DB	181	CTGGTGAAGCTGCTCTCTATGCAAGTCAGGGGCTGGGAGATGTTGTGGTTCGTTC	240		
QY	773	caatgctctgctcgtcgtggtccaccgagtgtgtgttaattgcaatgcaatgcaatg	832		
DB	241	CAATTGCTCTGCTGCTGCTGCTACCGGATGATGTTGTAATGCCAAGATACCTTAATG	300		
QY	833	ggtcctctgataaaactatgcaagggattatcacatgcgaagcacataagattccat	892		
DB	301	GGTCTCTGATAAAACTATGCAAGGCATATACATGCGAAGCACATTAAGATTCCAT	360		
QY	893	gctttggggatcacatgaagtgcaccttttctatgtagtagacacacgtcgtggg	952		
DB	361	GCTTTGGGGATCACATGAAGTGACCTTTTTCATGAGTATAGACACAACGTCGATTGGG	420		
QY	953	tgtttgcatcatccgtcatatcacagaccaggagtttatatgagataatttgggtg	1012		
DB	421	TGTTTTCGATCATCTGTCATATCAGACACAGGAAGTTTATATGGAGATAAATTTGGTG	480		
QY	1013	cttttgggtgataatcagttcagatacacactccttctgctatcgtcagcagggcccccac	1072		
DB	481	CTTTTGGTGATAATCAGTTTCAGATACACACTCCTTTTGCTATGCTGATCGAGGCCCCAC	540		
QY	1073	taatccttgaattgggaggataattatgacagaaattgcatgtttgttggaaacgatt	1132		
DB	541	TAATCCTTGAATTTGGGAGGATATATTATGGACAGAAATTCATGTTTGTGTGAACGATT	600		
QY	1133	ggcatgcagccttctgtgccagtccttctgtcgtcaaaatatagacatacgggtgtttaca	1192		
DB	601	GGCATGCAGCCTTGTGCCAGTCTCTTCTGCTGCAAAATATAGACCATACGGGTGTTTACA	660		
QY	1193	gagattcccgagcacaccctgttatatacataatttagcacatcaggtgtggagcctgcgaa	1252		
DB	661	GAGATTCCCGCAGCACCCCTTGTATTATATAAATTTAGCACATCAGGGTGGAGCCTGCAA	720		
QY	1253	gtacatactcgtgaggtggtcctcctggaatggtatggagcttttagaatgggtatttc	1312		
DB	721	GTACATATCCTGATCTGGGATTGCTCTGAAATGGTATGAGCTTTAGAAATGGGTATTC	780		

Qy	1313	cagaatgggcaagagagcagtcgccccttgacaaggggtgagcagtaactttttgaaagag	1372		
Db	781	CAGATGGGCAAGGAGGSCATGCCCCTTGACAAGGGTGAGGCACTTAACCTTTTGAAGAGG	840		
Qy	1373	cagtttgacagcagatcgagattgtacacgtcagtcagggttattcatgggggtcacaa	1432		
Db	841	CAGTTGTGACAGCAGATCGGATTGTGACCCCTCACTCAGGTTATTTCATGGAGGTCACAA	900		
Qy	1433	ctgctgaaggtggacaggggctcaatgagctcttaagctcccccacaaagtgtatgaaag	1492		
Db	901	CTGCTGAAGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAAGAGTGTATTGAATG	960		
Qy	1493	gaattgtaaatggaattgaaataatgattggaacccccacacagacaagtgtcctccctc	1552		
Db	961	GAATTGTAATGGAATTGACATTAAATGATTGGAACCCCAACACAGACAAAGTGTCTCCCTC	1020		
Qy	1553	atcattattctgcatgacacctctctgaaaggccaaatgtaaagctgaattgcagaag	1612		
Db	1021	ATCATATTCTGTCGATGACCTCTCTGAAAGGCCAAATGTAAAGCTGAAATTGAGAAGG	1080		
Qy	1613	agttgggtttaccctgtaaggagagatgttccctcgtgattggctttatttggaagactgatt	1672		
Db	1081	AGTTGGGTTTACCTGTAAGGAGGAGTGTCTCTGATTTGGCTTTTATTGGAGACTGGAAT	1140		
Qy	1673	accagaaaggcattgaatctcattaaaaaggccattccagagctcatgagggagacgtgc	1732		
Db	1141	ACCAGAAAGGCAATGATCTCATTAATGAGCCATTCAGAGCTCATGAGGAGGACGTGC	1200		
Qy	1733	aatttgcattgctggtgattggtggggtcccaatttttgaaggctggatgcagactaccaggt	1792		
Db	1201	AAATTGTCATGCTGTCGATCTGGGATCCAAATTTTGAAGGCTGGATGAGATCTACCCAGT	1260		
Qy	1793	cgagttacaagataaaatccgtggtggtggtggttgcagtttcagtttcccccacagaa	1852		
Db	1261	CGAGTTACAAAGGATAAATTTCCGTGGATGGGTGGATTAGTGTCCACTTTCCACAGAA	1320		
Qy	1853	taactcaggttgcgatatattgttaatgccatcgagattgaaacctgcgggtcttaate	1912		
Db	1321	TAATCTGAGTTGCGATATATTGTTAATGCCATCGAGATTTTGAACCTTTCGGGTCTTAATC	1380		
Qy	1913	agctatgctatgcaatatggtacagttcctgtagttcatggaactggggcctcccgag	1972		
Db	1381	AGCTATATGCTATGCAATATGTTGATACAGTTCTCTGATGTCATGGAACGGGGCCCTCCGAG	1440		
Qy	1973	acagtcagagaccttcaacccttttgggtgcaaaaggagaggggtacaggggtggggcgt	2032		
Db	1441	ACAGTCGAGACCTTCAACCCCTTTTGGTGCAAAAGAGAGAGGGGTACAGGGTGGGGGT	1500		
Qy	2033	tctcaccgctaaacgctggacaagatgttggggcattggaacccgagatgtcgacattca	2092		
Db	1501	TCTCACCGCTAACCGTGGACAAAGATGTTGTGGCATTTGGCAACCGCGATGTGCACATTC	1560		
Qy	2093	gggagacaagccgctcctgggaggggtcctgaagcgagggcagtgacgaaagaccatacgt	2152		
Db	1561	GGGAGCAAAAGCCGCTCTGGGAGGGGCTCATGAAGCGGCGATGACGAAAGACCATACGT	1620		
Qy	2153	gggacctgcccgcagcagtcacagcagatcttcagttggggccttcgtggacacaccccta	2212		
Db	1621	GGGACCATGCCCGCAGCAGTACGAGCAGATACCCCTGCTACATTTCGAGTCTCAAGACAT	1680		
Qy	2213	cgtcatgtagcggggactggggaggtccaaagtcgaagtcctcctcagctctgaagacat	2272		
Db	1681	CGTCATGTAGAGGGGACTGGGGAGGTCCAAAGTGCAGAGTCTCTCTTACAGTCTCAAGACAT	1740		
Qy	2273	cctctcactcctccggggcccggaaggaacccctctgacattgctgcttgcgtacag	2332		
Db	1741	CCTCTTCAATCTTCCGGCGGCCCGGAGGATACCCCTGCTACATTTCGTTGCTCTACAG	1800		
Qy	2333	tagatgcgaatgcgctcgttgccttgcgcgggttcgagacaacatacagcgctgtg	2392		
Db	1801	TAGAGTCGCAATGCGCCTGCTTGTGTTGCTGCGGGTTCGAGAACATATGACGCGCTGTG	1860		
Qy	2393	ctgctcggcggtgacagcttcgggttgacgacagttacagtttttgggggaaataaggaag	2452		

Db 390 TTGGAAGAGGAGCAGCTCGAGCTAAAGTAACACAAAAACATTGCTTTGTAACTGGCGAAG 449
QY 722 ctgctcctctatgcaaaagtcagggggttggaagatgtttgtgatttgcgttaccaattgctc 781
Db 450 CTCTCTCTATGNAAGCTCGGGGCTAGAGATGTTTGTGGTTTCATTGCCAGTGTGCTC 509
QY 782 ttgctgctgctggttcaccgagtgatggttgtaagcgaagatactactaaatgggttcctctg 841
Db 510 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 569
QY 842 ataaaaactatgcaaaagcattatacactcgaagcacattagaattccatgctttggggg 901
Db 570 ATAGAAATTAAGCAATGCAATTTACACAGAAAAACACATTCGATTCGATTCGATTCG 629
QY 902 gatcaatgaagtgcacctttttcatgtagtagagacaacgctgattgggtgtttgtcg 961
Db 630 GTGAACATGAAGTTACCTCTCTCCATGAGTATAGAGATTCAGTTGACTGGGTGTTTGTG 689
QY 962 atcatcgtctatcatcagaccagggaagtttatatgagataattttgtgcttttgggtg 1021
Db 690 ATCATCCCTCATATCAGACACCTGGAATTTATATGAGATAAGTTTGTGCTTTTGGTG 749
QY 1022 ataactcagttcagatacacactccttctgctatgctgcatcgagggcccaactaatccttg 1081
Db 750 ATATCAGTTTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTTGGTCTTG 809
QY 1082 aattggagagatatattatggacagaattgcatgtttgtgaaacgattgcatgccca 1141
Db 810 AATTGGGAGGATATATTATGGACAGATTCATGTTGTTGTCATGATGGCATGGCCA 869
QY 1142 gcttgtgcagtcctctctgctgcgaataatgaccatacagtggttttacagagattccc 1201
Db 870 GTCTAGTGGCAGTCTTCTTCTGCTGCAAAATATAGACATATGCTGTTTATAAAGACTCCC 929
QY 1202 gcagcacccttgtatatacaataattagcacatcaggtgtgagcctgcaagtacatatc 1261
Db 930 GCAGCATCTGTGTAATACATAATTTAGCACATCAGGCTGAGAGCTGCAGCACATATC 989
QY 1262 ctgactggtggttcctcctgaatgtagtgagcttttagaattgggtatttccagaatggg 1321
Db 990 CTGACCTTGGTTGCCACCTGAAATGATGATGGAGCTCGAGTGGGTATTCCCTGAATGGG 1049
QY 1322 caagggagagcctctgacaggggtgagggcaggttaactttttgaaaggagcaggtgtga 1381
Db 1050 CGAGGAGGATGCTTGTGCAAGGGGTGAGGAGCTTAATTTTGAAGGTGCAAGTGTGA 1109
QY 1382 cagcagatcgagttgtgacctgagtcaggggttattcatgggaggttcacaaactgctgaag 1441
Db 1110 CAGCAGATCGAATCGTGACTGTGCTAGAGGTTATTTCATGGGAGGTCAACAATGCTGAAG 1169
QY 1442 gtgacagggcctcaatgagctcttaagctcccgaaaaagtgtattggaatgggaattgtaa 1501
Db 1170 GTGGACAGGGCCCTCAATGAGCTTAAAGCTCCAGAAAGAGTGTTAAACGGGAATGTAA 1229
QY 1502 atggaattacattatgattggaaccccccacacagacaagtgctcctccatcattatt 1561
Db 1230 ATGGAATTGACATTAATGATTGGAACCTTCCACACACAAATGATCCCTGTGCATATT 1289
QY 1562 ctgtcgtgacacctctctggaagggccaaatgttaaagctgaattgcagaaggaggtgggtt 1621
Db 1290 CTGTTGATGACCTCTCTGGAAGGCCCAATGTAAAGGTCAATGTCAGAGGAGCTGGGT 1349
QY 1622 taccgtgaagggaggatgttcctctctgattggtcttatttgaagactggattaccagaag 1681
Db 1350 TACCTATAAGGCCCTGATGTCTCTGATTTGGCTTTATTGGAAGATTGGATTATCAGAAAG 1409
QY 1682 gcaattatctcataaaatggccattccagagctcatgagggagcgctgcgaatttgtaa 1741
Db 1410 GCATTGATCTCAATCAACTTATCATACCAAGATCTCATGCGGGGAAGATGTTCAATTTGTCA 1469
QY 1742 tgccttgatctctggggtcccaatttttgaagggctggatgagatctaccaggtcgagtaca 1801
Db 1470 TCGTGTGATCTGCTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCA 1529

QY 1802 aggataaaattccatggatgggttggtgatttagtttagttccagatttcccacagaataactgcag 1861
Db 1530 AGGATAAAATTTCTGGATGGTGGATTTAGTGTTCAGTTTCCACCGAATAACTGCCG 1589
QY 1862 gtgcgatatattgttaatgccatccagagatttgaacctgctgcttaatacagctatatg 1921
Db 1590 GCTGCGATATATTGTAATGCCATCCAGATTCCGAACCTTGTGCTCAATCAGCTATATG 1649
QY 1922 ctatgcaatattgtacagtttccctgttagtttcattggaactggggcctccagagacacagtcg 1981
Db 1650 CTATGAGTATGTCACAGTTCCTGTGTTCATCCAACTGGGGGCTTAGAGATACCGTGG 1709
QY 1982 agaccttcaaccttttgggtgcaaaaggagagaggttacaggttgggctttctcaccgc 2041
Db 1710 AGAATTTCAACCTTTTCGGTGGAATGGAGAGCAGGTTACAGGGTGGGCATTGCGACCCC 1769
QY 2042 taacctgtgacaagatgtgtggcattgcaaccgcgattgcgacattcagagagacaca 2101
Db 1770 TAACCACAGAAAACATGTTGTGCACATTGGGAATGCAATATCTACATACAGGGAACACA 1829
QY 2102 agcgttctctgggggggctcatgaagcaggcagtcagcaaaagaccatacagctgggacatg 2161
Db 1830 AGTCTCTCTGGGAAGGCTTAATGAAGCAGGATGTCAAAAGACTTTCACGTGGGACCATG 1889
QY 2162 cc-ccagcagctacgacagatcttcgagtgagccttcgtggaccaacctcagctcagtg 2220
Db 1890 CCCTGAACANTACGAACAAATCTTCCAGTGGGCTTTCATCGATCGACCCCTATGTCTATGT 1949
QY 2221 agacgggggactggggaggt 2239
Db 1950 AAAAAAGGACCAAAAGTGGT 1968

RESULT 9
AF036891
LOCUS
DEFINITION
Zea mays starch synthase I precursor (Ss1) mRNA, nuclear gene
ACCESSION
AF036891
VERSION
AF036891.1
KEYWORDS
GT:2828011
SOURCE
Zea mays.
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 2991)
AUTHORS
Knight,M.E., Harn,C., Lilley,C.E.R., Guan,H., Singletary,G.W.,
Mu-Forster,C., Wasserman,B.P. and Keeling,P.L.
TITLE
Molecular cloning of starch synthase I from maize (W64) endosperm
and expression in Escherichia coli
JOURNAL
Plant J. 14 (5), 613-622 (1998)
MEDLINE
98340555
PUBMED
9675904
REFERENCE
2 (bases 1 to 2991)
AUTHORS
Knight,M.E., Harn,C., Lilley,C.E.R., Guan,H.P., Singletary,G.W.,
Mu-Forster,C., Wasserman,B.P. and Keeling,P.L.
TITLE
Direct Submission
JOURNAL
Submitted (03-DEC-1997) ExSeed Genetics, 1568 Food Science
Building, ISU, Ames, IA 50011, USA
FEATURES
Location/Qualifiers
1..2991
/organism="Zea mays"
/strain="W64A"
/db_xref="taxon:4577"
/chromosome="9"
/map="9; between npi266 and wx1"
/tissue.type="endosperm"
1..2991
/gene="Ss1"
602..2470
/gene="Ss1"

[illegible]

QY 1296 tttagaatgggtatttccagaatgggcaagagcagtcgcttgacaaagggtaggcagt 1355
 Db 1075 TCTGGAGTGGTATTCCTTCAATGGCAAGAGGAGCATGCCCTTGACAAAGGTTGAGCAGT 1134
 QY 1356 taacttttgaaggagcagttgtgacagcagatcgattgtgacagtcagtcagtggtta 1415
 Db 1135 TAATTTTTTGAAGGTGCACTTGTGACAGCAGATCAATTTGTGACTGTAGTAAGGTTA 1194
 QY 1416 ttcattggaggtcacactctcaaggtgacagagccctcaatgagctcttaagctcccg 1475
 Db 1195 TTCATGGAGGTGACAACTGCTGAAGGTGACAGGGTCTCAATGAGCTCTTAAGCTCCCG 1254
 QY 1476 aaaaagtgtattgaatgaattgtaaatgaattgacattaatgattggaaaccccccac 1535
 Db 1255 AAAGAGTGTATTAACCGGAATGTAATGAATGACATTAAATGATTGGAACCTCGCAC 1314
 QY 1536 agcaagtgctccctcaatcatattctgtcgatgacctctctggaagggccaaatgtaa 1595
 Db 1315 GGACAAATGTATCCCTTGTCAATATTCTGTGTGATGACCTCTCTGGAAGGCCAAATGTAA 1374
 QY 1596 agctgaattcagaagagttgggtttacctgtaagggagagtggttccctctgattgctt 1655
 Db 1375 AAGTGCATTCGAAGAGAGCTGGGTTTACCTATAGGCCCTGAAGTTCCCTGTGATTGGTTT 1434
 QY 1656 tattggaagactggattaccagaaaagcattgatctcattaaatggccattccagagct 1715
 Db 1435 TATTGGAAGATTGGATTATCAGAAAGCATTTGATCTCATTTCAACTTATCATACCATCT 1494
 QY 1716 catgagggaagcagtcgaatttctcatgttgcattggtatctggggatcccaattttgaagctg 1775
 Db 1495 CATCGGGGAGCAGCTTCAATTTGTGATGCTTGGATCTGGTGACCCAGAGCTCGAAGACTG 1554
 QY 1776 gatgagatctaccagagtcgagtttacaagagataaattccgtggatgggttggatttagt 1835
 Db 1555 GATGAGATCTACAGAGTCGACTTCAAGGATAAATTTGCGTGGATGGGTTGATTTAGTGT 1614
 QY 1836 tccagtttccacagaaatactgcaggttgcgatatattgttaatgccatcgagattga 1895
 Db 1615 TCCAGTTTCCACCGGAATACTGCGGCTCGATATATTGTTAATGCAATCCAGATTCTGA 1674
 QY 1896 accttcggtcttaatacagctatagctgcaatattggtacagttcctgtagttcatg 1955
 Db 1675 ACCTTGTGTTCAATCAGCTATATGATGACAGTATGCGACAGCTTCTGTTGTCATGC 1734
 QY 1956 aactgggggctcgcagacacagtcgagaccttcaacccctttggtgcaaaagagagga 2015
 Db 1735 AACTGGGGGCTTAGAGATACTGTGGAGAACTTCAACCCCTTCGGTGAGAATGGAGAACA 1794
 QY 2016 ggtacaggggtggcggttctaccgctaaacggtgacaagatg-ttggggcattgcgaa 2074
 Db 1795 GGGTACAGGGTGGGCATTCGACCCCTTAACCCAGGAAACATGTTTGTGACATTTGCGAA 1854
 QY 2075 ccgcgatgcgacattcagggagcacaaagcgcctcctgggggggctcatgaagcgagga 2134
 Db 1855 CTGCAATTTCCGACATACAGAGCAGCAAAATCTTCTCGGAGGGCTCATGAAGAGGGCA 1914
 QY 2135 tgacgaagaccca tacytgggacctatgcc-ccgagcagtcagagcagatcttcagtgagg 2193
 Db 1915 TGTCAAAAGACTTCACCTGGGACCCTGCCGCTGAACAAATACGAACAATCTTCCAGTGGG 1974
 QY 2194 ccttcgtggaccacccatagctcatgtagcggggagctggggaggt 2239
 Db 1975 CCTTCATCGATCGACCCCTATGTGATGTAAGGGGGACCAACAGTGGT 2020

RESULT 11
 AR049919
 LOCUS AR049919 2990 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 1 from patent US 5824790.
 ACCESSION AR049919
 VERSION AR049919.1 GI:5971911
 KEYWORDS

SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 2990)
 AUTHORS Keeling, P. Lewis, Knight, M. E. and Guan, H.
 TITLE Modification of starch synthesis in plants
 JOURNAL Patent: US 5824790-A 1 20-OCT-1998;
 FEATURES Location/Qualifiers
 source l..2990
 BASE COUNT 758 a 655 c 801 g 776 t
 ORIGIN

Query Match 48.1%; Score 1332; DB 6; Length 2990;
 Best Local Similarity 81.7%; Pred. No. 3.8e-170;
 Matches 1617; Conservative 0; Mismatches 340; Indels 21; Gaps 6;

QY 237 cgaactctctccgctctctccctgtccgcgcgcacacacccatggcgcgacggcgct 296
 Db 537 CGCATTTGTTGCGCGGATCCACGGCCCCCCCCCGGCTCCTCTCTCTCTCTCTCT 596
 QY 297 cgcgcgcgggtgctcgcgcccgagctcgcctgcgcgcgacccgcgcgcgcgcgc 356
 Db 597 CCCAATGGGAGCCCTCGGCGCTGGGCGCCGCTGCCCTCTCTCTCTCTCTCTCT 656
 QY 357 ggcgtccgcctcgcctcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 407
 Db 657 GCGGCGCGCGCTCGCGAGACCGGGCGCGCGCGCGGAGGCTCCAGCGGTGTGTGCGCGCG 716
 QY 408 ctacgtcgcgcagctcagcagagggagggcgccgcgcgcgcgcgcgcgcgcgcgcgcgc 467
 Db 717 GTGCGTGGCGGAGCTGAGCAGGAGGGGCGCGCGCGCGCGCGCGCTGCCACCCCGCTGCT 776
 QY 468 ggcgcgcgcgcctcgtcgcagcttctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 527
 Db 777 GCGCGCGCGCTCGTCCCGGCTTCTCTCG--CGCGCGCGCGCGAGCCACCCAGGTCAGCC 833
 QY 528 ggc 587
 Db 834 GGCATCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGTCTCGA----- 887
 QY 588 gctcgaagggatgctcgcaggtattccatcgacagcataaattgctggtgcgaagtgcagcagga 647
 Db 888 ACCTGAAGGATTTGCTGAAGTTCCATCGATACACAGTAGTTGTGGCAAGTGCAGCAAGA 947
 QY 648 tctcagatcatgctgcgaatgagcaacctcaagcttaagttacacgttagcctcgtgct 707
 Db 948 TTCTGAGATTGTGTTGGAAGAGGAGGAAAGCTCGAGCTAAAGTAAACACAAAGCATTTGCTTT 1007
 QY 708 tctcagatcatgctgcctcctctatgc-aaagtcagggggggttgggagatgcttctggtggt 766
 Db 1008 TGTAAACCGCGAGCTTCTCTTAATCGAAAGTCTGGGGGCTCTAGGAGATGTTTGTGGTT 1067
 QY 767 cgttaccaattgctcgtcgtcgttcacgcagtgatggttgaatgtaacacgaact 826
 Db 1068 CATTCGACGTTGCTCTCTGCTCGCGCTCACCGTGTGATGTTGTAATGCCACAGCATTT 1127
 QY 827 taattgggtcctcctgataaaacatacgaaagcattatacacactgcgaagcacaataaga 886
 Db 1128 TAAATGTAACCTCCGATAAAGATTTATGCAAAATGCAATTTTACTCAGAAAAACACATTCGGA 1187
 QY 887 ttcactgcttgggggatacatgaagtgaccttttttctatgatgatagagacaacgtcg 946
 Db 1188 TTCCATTTCTTTGGCGGTGAACATGAAGTTTACCTTCTTCCATGAGTATAGAGATTCAGTTG 1247
 QY 947 attgggtgcttgcgatacctcgtcatcacagaccaggaagtttatatggagataatt 1006
 Db 1248 ACTGGGTGTTGTTGATCATCCCTCATATACAGACCTTGAATTTTATATGGAGATAAGT 1307
 QY 1007 ttggtgcttgggtgataaactcagtttcagatacacacactccttctgctatgctgcagtgagg 1066
 Db 1308 TTGGTGTCTTTTGGTGATAAATCAGTTTCAGATACACACTCCTTGTCTATGCTGCATGTGAGG 1367

QY 1067 cccactaactccttgaaatggaggatataatttatgacagaattgcattgttgtaga 1126
 Db 1368 CTCCTTTGGTCCCTGAAATGGGAGGATATATTTATGACAGAAATGCATGTTGGTGCA 1427
 QY 1127 acgatggcaatgcagcccttgccagctcctcttgctgcacaaatatagaccatacggtg 1186
 Db 1428 ATGATTGGCATGCGACGTAGACCCAGTCCCTTCTGCTGCAAAATATAGACCATATGGTG 1487
 QY 1187 ttacagagatccccagcaccccttggtatatacataatttagcacatcagggtgtaggc 1246
 Db 1488 TTTATAAGACTCCCGACGATTTCTTGTATACATAATTTAGCACATCAGGGTGTAGAGC 1547
 QY 1247 ctgcaagtacatactctgatctgggatgcctcctctgaattggtatgagcttttagaagg 1306
 Db 1548 CTGCAAGCACATATCTGACCTTGGGTGGCCACTGAATGGTATGAGCTCTGAGATGGG 1607
 QY 1307 tatttccagaatgggcaaggagcagtcctcttgacaagggtgaggttaacttttga 1366
 Db 1608 TATTCTCTGAATGGCGAGSAGCATGCCCTTGACAAGGTGAGGAGTTAATTTTTTGA 1667
 QY 1367 aaggagcagttgtacagcagatcgatigtgacctcagctcaggttattcatggagg 1426
 Db 1668 AAGTTCAGTTGTACAGCAGATCGAATCTGACTGTGAGTAAGGGTTATTCTATGGGAGG 1727
 QY 1427 tcacaactgctgaagtggtgacagggcctcaatgagctcttaagctcccgaaaaagtgtat 1486
 Db 1728 TCACAACCTGCTGAAGTGGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTAT 1787
 QY 1487 tgaatggaattgtaaatggaattgacataatgattggaaccccaacacagacaagtgc 1546
 Db 1788 TAAACGGAAATGTAAATGGAATTTGACATTAATGATTGGAACCTGCGACAGACAAATGTA 1847
 QY 1547 tcctcatcatattctgcagatcctctctgaaagcccaaatataaagctgaattgc 1606
 Db 1848 TCCCTTGCTATATCTGTGTGAGCTCTCTTGAAGGCTAAATGTAAGGTGTCATTGC 1907
 QY 1607 agaagagttgggtttaccttgtaaggagagtgcttctctgattggtctttatttggaagac 1666
 Db 1908 AGAAGAGCTGGGTTTACCTATTAAGCCCTGATGTTCTCTGTGTTGCTTTATTGGAAGAT 1967
 QY 1667 tggattacagaaagcattgatctcatataaatggccattccagagctcagtgaggag 1726
 Db 1968 TGGATTATCAGAAAGCATTTGATCTCATTCACATTAATACAGATCATCATCGGAAGA 2027
 QY 1727 acgtgcaattgtcatgcttgatctgggatcccaatttttgaagctggatgagatcta 1786
 Db 2028 ATGTTCAATTTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTA 2087
 QY 1787 ccgagtcgagttacaaaggataaattccgtggtatgggttgatttagttccagatttccc 1846
 Db 2088 CAGAGTCGATCTTCAAGGATAAATTTGCTGGATGGGTGGATTTAGTTCCAGTTTCCC 2147
 QY 1847 acagaataactgcaggttgcgatatattgttaatgcacagagattgaaacttgcggtc 1906
 Db 2148 ACCGAATAACTGGGGTGGCGATATATTGTTATGTCATCCAGATTCGAACCTTGTGGTC 2207
 QY 1907 ttaatcagctatatgtatgcaatattggtacagttcctgatttcatgaaactggggccc 1966
 Db 2208 TCNATCAGCTATATGCTATGACATATGGACAGTTCTGTTGTCATGCAACTGGGGGCC 2267
 QY 1967 tccgagacacagtcgagaccttcaaccccttttgggtcaaaaggagggaggttacagggt 2026
 Db 2268 TTAGAGATACCCGTGGAGAACTTCAACCCCTTTCGGTGAAGATGGAGAGGATACAGGT 2327
 QY 2027 gggcgcttcacccgcttaaccgttgagacaagatg-ttctgggcatggaaccgcgagtgcg 2085
 Db 2328 GGGCATTTCCACCCCTTAACACAGAAACATATGTTTGTGACATTTGCAACTGCAATATCT 2387
 QY 2086 acattcagggagcacaaagcctgctggagggtctcatgaagcagggcatgacaagac 2145
 Db 2388 ACATACAGGGACACAAGTAATTAATGGGAAGGGCTTAATGAAGCCAGGCATGTCAAAAGAG 2447

QY 2146 catacgtgggaccatgcc-ccgagcagatcagcagacagatcttcagtggtgctctgtg 2202
 Db 2448 TTCACGTGGGACCATGCCGTGACACATACAGCAAAATCTTCCAGTGGCGCTTCATCG 2505
 RESULT 12
 E06904
 LOCUS 2533 bp RNA PAT 29-SEP-1997
 DEFINITION CDNA encoding soluble starch synthase.
 E06904
 VERSION E06904.1 GI:2175060
 KEYWORDS JP 1994070779-A/1.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 2533)
 REFERENCE
 AUTHORS Baba,T. and Shimada,H..
 TITLE SOLUBLE RICE STARCH SYNTHETASE GENE AND ITS USE
 JOURNAL Patent: JP 1994070779-A 1 15-MAR-1994.
 MITSUI GIYOUSAI SHOKUBUTSU BIO KENKYUSHO:KK
 COMMENT
 OS Oryza sativa
 PN JP 1994070779-A/1
 PD 15-MAR-1994
 PF 07-JUL-1992 JP 1992179947
 PI BABA TADASHI, SHIMADA HIROAKI
 PC C12N15/52,A01H5/00;
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue=ripening seed;
 FH Key Location/Qualifiers
 FT 5'UTR 1..113
 FT CDS 114..1994
 FT /product='Soluble starch synthase' FT
 FT transit_peptide 114..452
 FT 3'UTR 1992..2533.
 FEATURES
 source
 1..2533
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 BASE COUNT 630 a 550 c 701 g 652 t
 ORIGIN
 Query Match 48.0%; Score 1329.4; DB 6; Length 2533;
 Best Local Similarity 80.7%; Pred. No. 8.8e-170;
 Matches 1606; Conservative 0; Mismatches 371; Indels 14; Gaps 4;
 QY 250 cctctccctccctgtcccgcgcgcacacccatggcggcagcggcgctgcgcgggtgc 309
 Db 54 CCGCCAAGCGCGCGCCGCCGACACAGCAGCAGCAGGATCGCGGAGAGAGGGGGATC 113
 QY 310 ctgcgccccagctgcgcctgcgcgcgatccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 369
 Db 114 ATGGCGACGGCGGGGGATGGGGATCGGGGGCGGCTGTGGTGGCGCGCAGGTGAGG 173
 QY 370 gtctgc 429
 Db 174 CCGGGGAGGAGGTTCGGCTCCAGCGGGTGGCGGAGGGGTGCGTGGCGGAGCTCAGCAGG 233
 QY 430 gaaggccccgc 483
 Db 234 GACGGTGGGTGCGCGACGCGCGCGCTGGCACCGCGCGCGCTGGTGAAGCAGCGGCTCTG 293
 QY 484 ccagggttctgc 537
 Db 294 CCGACCTTCTCTGTCGCGAGCTGACGCCACCGCGCGCGCTGCGCGCGCGCGCGCG 353
 QY 538 cagccgccctgc 597

Db 354 CCGACCCCGCGCGTTCGCGGATCTCGGCGTGGGGAGATCGAGCCGATCTAAGAGGT 413
 Qy 598 atgtctgaggtatccatcgacagacataatgttgctgctcaagctgagcaggatctcgagatc 657
 Db 414 CTCACAGAAGATCCATCGACAAACAATTTTGTGGCTAGTAGAGCAGGAGCTGAGATC 473
 Qy 658 atggatcggaatgagcaacctcaagctaaagtttacacgtagcactcggtttgtgactgggt 717
 Db 474 ATGGATGTGAAGGAGCAGAGCTCAAGCTAAAGTAAACGACGACGCTGTGTCTGTAACCGGT 533
 Qy 718 gaagctctctatgcaagtcagggggttggaagatgtttgttggttctggttaccat 777
 Db 534 GAAGCTTCTCTTATGCAAAATCAGGTGGACTAGGAGATGTTGTGTTCTCACTGCCAAT 593
 Qy 778 gctcttgctgctggtgcacccgagtgatggttgaatgccaagatacttaaatgggtcc 837
 Db 594 GCTCTGCTCTGCTGCTATCGTGTGATGCTGTTGTAATGCCGAGATACATCAACGGGGCC 653
 Qy 838 tctgataaaactatgcaagggaattacactgcgaagcacataagattccatgcttt 897
 Db 654 TTGAACAAAAATTTTGGAAACGCACTTTTACACTGAGAGACACATTAAGATTCCATGCTTT 713
 Qy 898 ggggatcacatgaagtcgactcttttcatgagtatagacacaacgtcgattgggtttt 957
 Db 714 GCGGAGAACATGAAGTTACTTTTTTTCACAGATATAGGATCTGTGTGATGGGTGTTT 773
 Qy 958 gtcgatactcgtcatatcacagaccaggaagtttatagagataattttgggtctttt 1017
 Db 774 GTTGATCATCCCTCATATCATAGACCTGGAAATTTGTATGGAGATAATTTTGGTCTTTT 833
 Qy 1018 ggtgataatcagttcagatacaactccttctgctatgctgcgagggcccactaatc 1077
 Db 834 GCGGATAATCAGTTTTCAGATACACACTCTGCTGTATGCGCGTGTGAAGCCCGATTAAT 893
 Qy 1078 ctggaattggagagatatttatggacagaattgcattgtttgttgaacgattgcat 1137
 Db 894 CTGGAACCTGGGAGGATATCTATGGACAGAAATGCATGTTTGTGTGAATGATGGCAT 953
 Qy 1138 gccagcctgtgcccagtccttcttctgctgcaaaatatagaccatacgcgtgttttacagat 1197
 Db 954 GCCAGTCTGTGCCAGTCCCTTCTTGTCTGCAAAATATAGACCATATGTTGTACAGGAT 1013
 Qy 1198 tcccgaagacccttggtatacataaatttagcacatcagggtgtggagcctgcaagtaca 1257
 Db 1014 GCGCGCAGTCTTCTGTATACATAAATCTAGCACATCAGGCTGTGGAGCCTGCCAGTACA 1073
 Qy 1258 tatectgctgggattgctcctggaatgtagtgagctttagaattggtatttccagaa 1317
 Db 1074 TATCTGACCTGGGATGCGCACCTGAATGATGAGACATTAGAATGGGTGTTTCCAGAG 1133
 Qy 1318 tgggcaaggaggatgcctctgcaagggtgagcgagtttaacttttgaaggagcgagtt 1377
 Db 1134 TGGCAAGGCGCATGCTCTTGCAAGGCTGAGCAGTCAATTTTAAAGGCGCAGTT 1193
 Qy 1378 gtacagacagatcggaattgtgacctgagtcaggttatttcaggggttcaacaactgct 1437
 Db 1194 GTACAGCAGATCGAATTTGTGACTGTGACGACAGGGGTATTCTATGGGAGGTACAACTGCT 1253
 Qy 1438 gaaggtgacagggcctcaatgagctcttaagctcccgaaaagtgtatgaattggaatt 1497
 Db 1254 GAAGTGGGCAAGGCCCTCAATGAGCTCTTAAGCTCCCGGAGAGGTGATTGAATGGAAT 1313
 Qy 1498 gtaaatggaattgacataatgattggaacccaccacagacagagtgctctccctcatcat 1557
 Db 1314 GTRAAATGGAATTGACATTATGATTGGAACCCATCCACAGACAGTTTCTCCCTTATCAT 1373
 Qy 1558 tattctgcatgacacctctctggaagggccaaaatgtaaaactgaattgcagaaggagttg 1617
 Db 1374 TATTCTGTTGATGACCTGTCCGGAAGGCCAAAGTGTAAAGTGAATTTGCAGAGGAGCTG 1433
 Qy 1618 ggttacctgaaggaggaggttctcctgattggtctttttagaacgaggtattaccag 1677

Db 1434 GGTTTACTATAGGCCCGATGTGCCCTCTGATTCGGCTTTATTGGAAAGATTGGACTATCAA 1493
 Qy 1678 aaagcattgatctcattaaaatggccattccagagctcagtgaggagcgctgcaattt 1737
 Db 1494 AAAGCATTTGATCTAATTAACCTTGGCCATCCAGATCTCATGCGGACAAATATTCAATTC 1553
 Qy 1738 gtcagtgttgatctggggatcccaatttttgaaggctggatgagatctaccagagtcaggt 1797
 Db 1554 GTCATGTTGGATCTGGTGCACCCAGGTTTGAAGGATGGATGAGATCCACAGAATCAGGG 1613
 Qy 1798 tacaagataaaatccgtgcatggttggtttagtttagttccagttccccacagataaact 1857
 Db 1614 TACAGGATAAATTTTCGTGGATGGTTGGATTTAGTGTTCAGTTTCCCACCGAATRACT 1673
 Qy 1858 gcaagttgcgataatattgtaatacccatcagagatttgaaccttcggtcttaatacagcta 1917
 Db 1674 GCAGTTTGCAGATATATTGTTGATGCCATCCAGATTTCGAACCTTGTGCTCAATCAGCTA 1733
 Qy 1918 tatgctatgcaatattggtacagtttccctgtagtttcatggaacctggggcctccggagaaca 1977
 Db 1734 TATGCTATGCAATATGTTACAGTGCCTGTTGTTTCATGGAATTCGGATTCAGAGATACA 1793
 Qy 1978 gtcagagacctcaacctcttgggtgcaaaagagagaggtacagaggtggcgcttctca 2037
 Db 1794 GTGAGAAATTTTAAACCGTTTGTCTGAGAAAGAGAGCAGGTACAGGGTGGCATTCG 1853
 Qy 2038 ccgctaacc-gtggacaagattgttggcatttgcgaacccgcgagatctgcacattcagga 2096
 Db 1854 CCATAACCATTTGAAAAAATGCTGTGGCATTCGGATTCGGATTCGCAATTCGACATACAGGA 1913
 Qy 2097 gcacaaagcctcctggaggggctcatgaaagcagaggtacgaaagaccatacagtgaggga 2156
 Db 1914 ACACAACTCCTCTTGGGAGGCTCTAATGAAGCAGGAGATGTCGAAGCAGCTTTTACATGGGA 1973
 Qy 2157 ccattgce-ccgagcagcagcagcagcttctgagtgaggttcctgagcaacccctcag 2215
 Db 1974 CCATCGCGCTCACAGTATGACAGATCTTCGAATGGCCCTTCATGATCAACCATATGT 2033
 Qy 2216 catgtagacgg 2226
 Db 2034 CATGTAATGG 2044

RESULT 13

RICSSS 2533 bp mRNA PLN 01-FEB-2000
 LOCUS Rice mRNA for soluble starch synthase, complete cds.
 DEFINITION
 ACCESSION D16202
 VERSION D16202.1 GI:450484
 KEYWORDS soluble starch synthase.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1 (bases 1 to 2533)
 AUTHORS Baba,T., Nishihara,M., Mizuno,K., Kawasaki,T., Shimada,H., Kobayashi,E., Ohnishi,S., Tanaka,K. and Arai,Y.
 TITLE Identification, cDNA cloning, and gene expression of soluble starch synthase in rice (Oryza sativa L.) immature seeds
 JOURNAL Plant Physiol. 103 (2), 565-573 (1993)
 MEDLINE 94302151
 REFERENCE 2 (bases 1 to 2533)
 AUTHORS Baba,T.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-1993) to the DDBJ/EMBL/GenBank databases. Tadashi Baba, University of Tsukuba, Institute of Applied Biochemistry; Tennohdai 1-1-1, Tsukuba Science City, Ibaraki 305, Japan (Tel:298-53-6632, Fax:298-53-6632)
 COMMENT Submitted (06-May-1993) to DDBJ by: Tadashi Baba
 Institute of Applied Biochemistry
 University of Tsukuba

Qy	1198	tccgcagcagcccttggtatatacataattttagcacatcaggggtgtggagccttcgaagtaca	1257
Db	1039	GCOCGAGTCTTCTTCATACATATATCTAGACACATCAGGGTGTGGAGCCTTCGCACGTACA	1098
Qy	1258	tatcctgatctgggattgctcctcgaatgtgatggacttttagaatgggtatttccagaag	1317
Db	1099	TATCCTGACCTGGGATTGCCACCTGNAATGGTATGGAGCATTTAGAAATGGGTCTTTCACAGAG	1158
Qy	1318	tgggcaaggaggcatgcctccttgacaagggtgaggcagttaactttttgaaaaggagcagtt	1377
Db	1159	TGGCAAGGCGGATGCCCTTGACAAAGGTGAGGCAGTCAATTTTAAAAAGGCGCAGTT	1218
Qy	1378	gtgacagcagatcggtattgtaccctcagtcaggttattcatcgtggaggtcacaactgct	1437
Db	1219	GTGACACAGATCCAAATGTGTACTGTACGCCAGGGGTATTATCGGAGGTCAACAACGTCT	1278
Qy	1438	gaagggtggacagggtccatgaagctcttaagctcccgaaaaaagtgtattgaatggaaatt	1497
Db	1279	GAAGGTGGGCAAGGCCATCAATGAGCTCTTAAGCTCCCGAAGAGTGTATTGAATGGAATT	1338
Qy	1498	gtaaatggaattgacaattaatgatgtgaaacccccacacagacagaagtgctcctcctcatc	1557
Db	1339	GTAAATGGAAATGACAAATTAATGATTGGAAACCCATCCACAGACAAGTTTCTCCCTTATCAT	1398
Qy	1558	tattctgtcgatgacctctctgtgaaaggccaaatgtaaagctgaattgcagaagaggttg	1617
Db	1399	TATTCGTTGATGACCTGTCGCGAAGGCCAAGTGTAAAGCTGAAATGTCAGAACGACGTG	1458
Qy	1618	ggttacctgtgaaggaggatgtcctcctgtatggctcttatgtgaaagactgattaccag	1677
Db	1459	GGTTTACCTATTAAGGCCGATGTGCTCTGATTGTGCTTTATTGGAAGATTGGACTATCAA	1518
Qy	1678	aaaggcatgatctcattaaatggccattccagagctcatgaggagagcgtcacaattt	1737
Db	1519	AAAGGCATTTGATCTAATTAACCTTGCCATTCCAGATCTCATGCGGGACAAATATTCAATTC	1578
Qy	1738	gtcatgcttbgatctggggatcccaatttttgaaggctggatgagatctaccgagtcgagt	1797
Db	1579	GTATGCTTGGATCTGGTCACCCAGGTTTGAAGGATGGATGAGATCCACAGAATCAGGG	1638
Qy	1798	tacaagataaaatccgtggatgggttgatttagtcttcaggttccccacagaataact	1857
Db	1639	TACAGGGATAAATTTCTGCGATGGGTGGATTTAGTGTCCAGTTTCCACCGAATAACT	1698
Qy	1858	gcagggttcgataatgtttaatgcatacgagattgaaaccttgcggtctttaaactcagcta	1917
Db	1699	GCAGGTGCGATATATTGTTGATGCCATCCAGATTCGAACCTTGTGGCCTCAATCAGCTA	1758
Qy	1918	tatgctatgcaatatggtacagtctcctgtagtctatgaaactgggggctccgagacaca	1977
Db	1759	TATGCTATGCAATATGGTACAGTGCCTGTTGTTCATGGAACCTGAGGCCTCAGAGATACA	1818
Qy	1978	gtcgagaccttcaacccttttggtcaaaagagagaggtacaggttgggcgcttctca	2037
Db	1819	GTGGAGAAATTTAACCCGTTTCTGTAGAAAGGAGACAGGTACAGGGTGGGCATCTCTCG	1878
Qy	2038	ccgctaacc-gtggacaagatgtgtgggcattgcgaaccgcgatgtcgacattcaggga	2096
Db	1879	CCACTAACCAATTTGAAAAAATGCTGTGGCATTTGCGGATTGCAATTTTCGACATACAGGA	1938
Qy	2097	gcacaagcgtctcggagggggctcatgaagcgaagcatgacgaaaagacatacgtggga	2156
Db	1939	ACACAAGTCTCTTTGGAGGGGTCTAATGAAGCGAGGCATGTCAAGCGACTTTTACATGGGA	1998
Qy	2157	ccatgcc-ccgagcagtacgacagatcttcgagtgggccttcgttgaccaaaccttcgct	2215
Db	1999	CCATGCCGCTTCACATGTGAACAGATCTTGAATGGGCCCTTCATGGATCAACCATATGT	2058
Qy	2216	catgtagaagg 2226	
Db	2059	CATGTAATGG 2069	

RESULT	15	A63308	2992 bp	DNA	PAT	12-MAR-1998
LOCUS		A63308	2992 bp	DNA		
DEFINITION		Sequence 1 from Patent WO9720936.				
ACCESSION		A63308				
VERSION		A63308.1	GI:3717138			
KEYWORDS		unidentified.				
SOURCE		unidentified.				
ORGANISM		unclassified.				
REFERENCE		1 (bases 1 to 2992)				
AUTHORS		Keeling, P.L. and Knight, M.E.				
TITLE		MODIFICATION OF STARCH SYNTHESIS IN PLANTS				
JOURNAL		Patent: WO 9720936-A 1 12-JUN-1997;				
		ZENECA LTD (GB)				
COMMENT		Other publication AU 1037197 19970627.				
FEATURES		Location/Qualifiers				
source		1..2992				
		/organism="unidentified"				
		/db_xref="taxon:32644"				
		/clone="NUMBER 1"				
BASE COUNT		758 a 655 c 801 g 776 t				
ORIGIN		2 others				
Query Match		47.9%; Score 1327.6; DB 6; Length 2992;				
Best Local Similarity		82.0%; Pred. No. 1.5e-169;				
Matches 1606; Conservative		0; Mismatches 335; Indels 18; Gaps				
QY	247	ccgcctctccctgtcccgcgccacacccatgcgcgacggcggtcggcgcggg	306			
Db	563	ccccccccccgcgtctgtctgtcttcccttccgcaatggcgacccctcgccgctg	622			
QY	307	tgctcgccccagctgcgcgcctgcgcgcgatccgcgacggcgccggcgctccgc	366			
Db	623	ggcgccgctgctctctctcgccgggncgcttgccgcccgcgcgtcg--CGAC	676			
QY	367	tgctgttcgcgcgcgcgtcgcgcgttggcgggcgccctaagtcgcgcagctcagc	426			
Db	677	cgggcgcccgccggagagctcagcgctgtgtgcgcgcgcgtgcgtcgcgagctgac	736			
QY	427	aggagggcccgcgcgccgcgcgcgcagcagcgaactggcccgccgctcgtgcga	486			
Db	737	aggagggcccgccgcgcgcgcgcgtgccaccgcgcgtgtggcgcccgccgctgcgc	796			
QY	487	ggcttcctgc	546			
Db	797	ggcttctcgc	853			
QY	547	ctgcgacgc	606			
Db	854	gtgcccgc	907			
QY	607	gattccatcgacagcataattgtgctgcaagtgcgcgcgcgcgcgcgcgcgcgcgc	666			
Db	908	ggttccatcgatcacacagtagttgtgcgaagtgcgcgcgcgcgcgcgcgcgcgcgc	967			
QY	667	aatgac	726			
Db	968	aagagcgaagctcgagctttaaagtagacacacacacacacacacacacacacacac	1027			
QY	727	ccctt-atcaagtcagggggttggagagatttgggttcgttaccaattgctcttgc	785			
Db	1028	ccttaatcgaaagctcgggggtcaggagatgttgggttcgttaccaattgctcttgc	1087			
QY	786	tgctgtggtcacagtgatggttgaatgccaagataactaaatgggtcctctgataa	845			
Db	1088	tgctgc	1147			
QY	846	aaactatgcaaggaattatcacctgcggaagcacataagattccttgggggac	905			
Db	1148	gaattatgcaatgcatTTTTTACTCAGAAAAACACATTCGATTCCTTTTGGCGTGA	1207			

Qy 906 acatgaagtgcaccttttttcatgagtagagagacaacgctcgattgggtgttggctgatca 965
Db 1208 ACATGAAGTTACCTCTTCCATAGTATAGAGATTCACTGGGTGTTGTTGATCA 1267
Qy 966 tccgtcatacacagacagaaagtcttatggagataaatttggctgttggctgataa 1025
Db 1268 TCCCTCATATCACAGACCTGGAATTTATATGGAGATTAAGTTGGTGTCTTTGGTGATAA 1327
Qy 1026 tcagttcagatacacactcctttgtctatgctgcagcgagcccccactaataccttgaatt 1085
Db 1328 TCAGTTTCAGATACACACTCCTTTCTCTATGCTGTCATGTGAGGCTCCTTTGGTCTGAATT 1367
Qy 1086 gggaggatataattatgagacagaattgcattgttggctgtaacagattggaatgcagacct 1145
Db 1388 GGGAGGATATATTTATGACAGAAATTGCATGTTGGTGTCTCAATGATTGGCATGCCAGTCT 1447
Qy 1146 tgtgccagtccttctgctcaaaatagaccatacagtggttttacagagattcccgag 1205
Db 1448 AGACCAAGTCTCTTCTGCTCAAAATATAGCCATATGGTGTATTAAAGACTCCCCGAG 1507
Qy 1206 cacccttgttatacataatttagcacatcaggggtgtgagcctgcaagtagacatatcctga 1265
Db 1508 CATTCCTGTAATACATAAATTAGCACATCAGGGGTAGAGCCTGCAAGCACATATCCTGA 1567
Qy 1266 tctgggattgcctcctcctgaattggtatgagatttagaattgggtattccagaatgggcaag 1325
Db 1568 CTTTGGGTTGCCACCTGAATGGTATGGAGCTTGGAGTGGGTATTCCTGAAATGGGCGAG 1627
Qy 1326 gaggcatgcccctgacaggggtgagcgagtttaactttttgaaagagcagttgtgacagc 1385
Db 1628 GAGGCATGCCCTTGACAAAGGCTGAGGCAGTTAATTTTTGAAAGTGCAGTTGTGACAGC 1687
Qy 1386 agatcggattgtgacctcagtcaggggttattcaatgggaggtcacaaactcgtgaaggtg 1445
Db 1688 AGATCGAATCGTGACTGTCAGTAAGGGTTATTCTAGGAGTGCACAACTGCTGAAGGTGG 1747
Qy 1446 acagggcctcaatgagctcttaagctcccgaaagtgattgaattgtaaatgg 1505
Db 1748 ACAGGGCTCAATGAGTCTTAAGCTCCAGAGAGGTATTAACCGGAATTTGTAATGG 1807
Qy 1506 aattgacattgaattggaacccaccacagacagaagtgctcctccctcatcattctgt 1565
Db 1808 AATTGACATTAATGATTGGAAACCTGCCACAGACAAATGTATCCCTGTCATATTCTGT 1867
Qy 1566 cga tgcacctctctggaagcccaaatgaaagctgaattcagaagaggttgggtttacc 1625
Db 1868 TGATGACCTCTCTTGAAGGCTAAATGTAAAGGTGCATTCAGAAAGGAGCTGGGTTTACC 1927
Qy 1626 tgtaaggagagatttctcctgattggtcttattggaagactggattaccagaaggcat 1685
Db 1928 TATAAGGCTGTATGTTCTGTATTGGTGTATTGGAGATTTGGATTATCAGAAAGGCAT 1987
Qy 1686 tgatctcatlaaaatggccattccagagctcagtaggagagacgtgcaatttgcattgct 1745
Db 1988 TGATCTCATTCACATTATCATACCAGATCTCATCGGAAGAAATGTTCAATTTGTCATGCT 2047
Qy 1746 tggattctggggatcccaatttttgaaggctggatgagatccacaggtcgagttacaagga 1805
Db 2048 TGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGA 2107
Qy 1806 taaattccgtgagattggtgattagttccagtttccacagagaataaactgcaggttg 1865
Db 2108 TAAATTTCTGGTGGGTTGGATTTAGTGTTCAGTTTCCACCGAATAACTGGCGCTGG 2167
Qy 1866 cgatatattgttaatgccaatgagatttgaaccttgcggtcttaattcagctatatgctat 1925
Db 2168 CGATATATTGTTAATGCCATCCAGATTGCAACCTTGTGGTCTCAATCAGCTATATGCTAT 2227
Qy 1926 gcaataatggtacagttcctgtgagttcatgaaactgggggctccgagacagtcagagac 1985
Db 2228 GCAGTATGGCACAGTTCCTGTTGTCCTGCAACTGCGGGGCTTAGAGATACCGTGGAGAA 2287
Qy 1986 cttcaaccttttgggtgcaaaaggagaggggtacaggggtggcggttctcaccgcgtaac 2045

Db 2288 CTTCAACCCCTTTCGGTGAGAATGGAGAGACAGGTTACAGGGTGGGCATTCGACCCCTAAC 2347
Qy 2046 cgtgacaagaatg-ttgtgggcatttgcgaaccgcgattgtcgacatttcaggagagacaaagc 2104
Db 2348 CACAGAAAACATGTTTGTGGACATTCGAACTGCAATATCTACATACAGGGAACACAAGT 2407
Qy 2105 cgtcctggaggggctcatgaagcagagcatgacgaaagacca tacytgggaccatgcc- 2163
Db 2408 AATAATGGGAAGGGCTAATGAAGCAGGCATGTCAAAAGAGTTCCAGTGGGACCATGCCG 2467
Qy 2164 ccgagcagttacgagcagatcttctgagttggccttcgtgg 2202
Db 2468 CTGAACAATACGAACAAATCTTCAGTGGGCTTCATCG 2506

Search completed: March 28, 2002, 14:54:59
Job time: 9898 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Db	1861	ctgctcgccggtgacagcttcggtggaacagcttacagtttggggaataaggaagg	1920
Qy	2453	gatgtctgcagatggtttaacagacaagcaccactcagatggcagcctctctgtccggtg	2512
Db	1921	gatgtctgcagatggtttaacagacaagcaccactcagatggcagcctctctgtccggtg	1980
Qy	2513	ttacagctgaatcagaacaccactggtgactcttttagccttagtgattgtgaaagtgtgt	2572
Db	1981	ttacagctgaatcagaacaccactggtgactcttttagccttagtgattgtgaaagtgtgt	2040
Qy	2573	tgcctctctgtatgtgtgtctgtccttagctgacaaatatttgacctgttggagaattt	2632
Db	2041	tgcctctctgtatgtgtctgtccttagctgacaaatatttgacctgttggagaattt	2100
Qy	2633	tatctttgtctgtgtttttttaatacaaaagaggggttctctccgatattcataaaaa	2692
Db	2101	tatctttgtctgtgtttttttaatacaaaagaggggttctctccgatattcataaaaa	2160
Qy	2693	aaa	2752
Db	2161	aaa	2220
Qy	2753	aaaaaaaaaaaaaaaaaaaaa 2771	
Db	2221	aaaaaaaaaaaaaaaaaaaaa 2239	
RESULT 4			
AAT95785			
ID	AAT95785 standard; cDNA to mRNA; 2383 BP.		
XX			
AC	AAT95785;		
XX			
DT	22-MAY-1998 (first entry)		
XX			
DE	Maize starch synthase type I cDNA.		
XX			
KW	Maize; starch synthase type I; starch; ds.		
XX			
OS	Zea mays.		
XX			
FH	Location/Qualifiers		
FT	2..1951		
FT	CDS		
FT	/*tag= a		
FT	/product= starch_synthase_type_I		
XX			
PN	DE19619918-Al.		
XX			
PD	20-NOV-1997.		
XX			
PF	17-MAY-1996; 96DE-1019918.		
XX			
PR	17-MAY-1996; 96DE-1019918.		
XX			
PA	(PLAN-) PLANTTEC BIOLOGIE GMBH.		
XX			
PI	Frohberg C, Kossmann J;		
XX			
DR	WPI; 1998-000821/01.		
DR	P-PSDB; AAW38218.		
XX			
PT	DNA encoding maize starch synthase type I protein - for producing		
PT	transgenic plants		
XX			
PS	Claim 1; Pages 16-20; 23pp; German.		
XX			
CC	The present sequence encodes maize starch synthase type I, useful		
CC	in the production of starch. Starch can be used in various		
CC	conventional starch applications, e.g. starch hydrolysate products,		
CC	foods, papermaking, adhesives, textiles, building materials, soil		
CC	stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal		
CC	briquettes, ore and coal slurries, foundry casting, rubber, leather		

THIS PAGE BLANK (USPTO)

```

Db 1861 CTGCTGCGCGGTGACACCTTCGGGTGGACGACAGTTACAGTCTTGGGAATAAGGAAGG 1920
Qy 2453 gatgtgctgaggtggttaacagcaagcaccactcagatggcagcctctctgtccgtg 2512
Db 1921 GATGTGTCAGGATGTTAAACAGCAAGCACCACACTCAGATGGCAGCCTCTGTGTCGGTG 1980
Qy 2513 ttacagctgaatcagaaacaaactggtgactctttgactctttgactctttgactctttg 2572
Db 1981 TTACAGCTGAATCAGAAACAACTGGTGACTCTTTAGCCTTAGCTAGTATTGTAAGTTGT 2040
Qy 2573 tgccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2632
Db 2041 TGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2692
Qy 2633 tatcttctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2752
Db 2101 TATCTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2812
Qy 2693 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2872
Db 2161 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2932
Qy 2753 aaaaaaaaaaaaaaaaaaaaaa 2771
Db 2221 AAAAAAAAAAAAAAAAAAAAAA 2239

RESULT 7
TAU48227
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
Source

CDS
/organism="Triticum aestivum"
/cultivar="T.a., Florida"
/db_xref="taxon:4365"
/tissue_type="endosperm of kernels"
<L. 1834
/codon_start=2
/product="soluble starch synthase"
/protein_id="AB02197.1"
/db_xref="GI:1373150"
/translation="EAPYAKSGGLGVDCGLPALTALARGHVRVWVMPRYLNGSSDKN
YAKALYTAHKIKPCFGSGHEVTFHEYRDNDVDFVDPHSYHRPGLSDGNFGAFGD
NOFRYTLICYAACEAPLILEGGYIYGONCMFVNDWHASLVPLVLAARYPYGVYRD
AVNTADRVTVSQYSWEVTTAEGLNELLSSKSVLNGVINGVINDNDNPTDKC
LPHYSVDDLSGAKAEQKELGLPREDVPLIGFGLRDYQKGLDLKMAIPELM
EPDGNOLYAMQYCTVPVHCTGRLDVTETFPNFGKGEETGWAFSPLTVDKMLWA
LRTAMSTFHEKPSMEGLMKRGHDKDTHWAPSSTSRSSGSPSWPTSCRLGLRS
KCESPSALKYSSSFRPEGPCTLRCPATVESQACLLWFAGSRVTDGCAAAAVTAS
GGROLPWGIKKAAGLWAKHSDGSLSVRTAEIRNLVL"
BASE COUNT 570 a 399 c 534 g 552 t
ORIGIN

```

```

Query Match 74.1%; Score 2054; DB 8; Length 2055;
Best Local Similarity 100.0%; Pred. No. 3.7e-267;
Matches 2054; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 718 gaagctgctccttatgcaaaagtcagggggttgaggagatgtttgttggttaccgaatt 777
Db 2 GAAGCTGCTCTTATGCAAAAGTCAGGGGGTGGGAGATGTTGTGCTGCTTACCAATT 61
Qy 778 gctcttgcctcgtggtcacccagtgatggttgtaatagcgaagatacttaaatgggtcc 837
Db 62 GCTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy 838 tctgataaaactatgcaaaagcattatcacctgcgaagcatttaagattccatgcttt 897
Db 122 TCTGATAAAACTATGCAAAAGCATTTATACACTGCGAAGCAGATTAAAGATTCCATCTTT 181
Qy 898 ggggagatcacatgaagtgaccttttttcatgagtagatagagacaaagtcgattgggttt 957
Db 182 GGGGATCACATGAAGTGACCTTTTTCATGAGTATAGAGACACGTCGATTGGGTGTTT 241
Qy 958 gtcgataatcgttcagatacacactcctttgtatgagtagatagagacaaagtcgattgggttt 1017
Db 242 GTCGATCATCCGTCATATCACAGACAGGAGCTTTTATATGAGATATATTTGGTGTGTTT 301
Qy 1018 ggtgataatcagttcagatacacactcctttgtatgagtagatagagacaaagtcgattgggttt 1077
Db 302 GGTGATAATCAGTTCAGATACACACTCTCTTGTATGCTGCTGCGAGGCCACCTAATC 361
Qy 1078 ctgaaatgggagatataatttggacagaaatgacacactcctttgtatgagtagatagagacaa 1137
Db 362 CTTCAATTTGGGAGGATATATTTATGGACAGAAATTTGATGTTGTTGTTGTTGTTGTTGTT 421
Qy 1138 gcaagccttgcagcctcctttgtgcaaaatagacacacactcctttgtatgagtagatagagacaa 1197
Db 422 GCCAGCCTTGTGCGAGCTCTTCTTGTGCAAAATATAGACCATACGTTGTTTACAGAGAT 481
Qy 1198 tcccgagcacccttcttataataatttagacacactcaggggtgagccttgcgaagta 1257
Db 482 TCCGCGACGACCTTGTATACATAATTTAGACATCAGGGTGTGGAGCCTGCAAGTACA 541
Qy 1258 tctctgctggtggtgctcctcctggaatggtgagcttgaatggtggttattccagaa 1317
Db 542 TATCCTGATCTGGGATTCCTCTCTGATGCTATGGAGCTTTTGAATGGGTATTTTCCAGAA 601
Qy 1318 tgggcaagagcattcctctgcaaggggtgagggcagtttaacttttgaagagcagtt 1377
Db 602 TGGCAGAGGAGGATGCTTCAAGAGGTGAGGAGTAACTTTTGAAGAGGAGGAGT 661
Qy 1378 gtgacagcagatcggaattgtgacccgtcagtgaggttattcattgaggtgcacactgct 1437
Db 662 GTGACAGCAGATCGGATTCGCGCTGCTGAGGGTATTCATGGAGGTCACAACTGCT 721
Qy 1438 gaagtgagcagggcctcaatgagctcttaagctcccccagcaagtgctcctcctcat 1497
Db 722 GAAGTGGACAGGGCTCAATCAGCTCTTAACTCCCGAAAAGGTATTTGAATGGAAT 781
Qy 1498 gtaaatgaattgacatttaattggaagcccaaaatgaagctgaattgagaggttg 1557
Db 782 GTAATGGAATGACATTAATGATTTGAACCCACACAGACAACTGCTCCTCATCAT 841
Qy 1558 tctctgctgagctcctctctggaagcccaaaatgaagctgaattgagaggttg 1617
Db 842 TATCTGTCGATGACCTCTCTGGAAGGCCAAATGTAAGCTGAATTGAGAGGAGTGG 901
Qy 1618 gtttacctgtaagggaggtgttctctgattggtcttattggaagctgaggtaccag 1677
Db 902 GGTTTACTGTAAAGGAGGATGTTCTCTCTGATTGGCTTTATTTGGAAGACTGGATTACAG 961
Qy 1678 aaaggcattgattcattaaatggccattccagagctcatgagggagcagtcgaattt 1737
Db 962 AAAGGCATTGATCTCATTAATAATGGCCATTCAGAGCTCATGGGAGGAGGAGGAGTCAATTT 1021
Qy 1738 gtcagctgtggtctgggggtcccaatttttgaagggctgagatgagatctaccagtgagt 1797

```

THIS PAGE BLANK (USPTO)

2042 AAAAAAAAAA 2055

U

[illegible]

Query Match 50.3%; Score 1393; DB 6; Length 2383;
Best Local Similarity 83.5%;
Matches 1619; Conservative Pred. No. 2.5e-178;

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2002, 12:57:36 ; Search time 249.65 seconds
(without alignments)
9515.915 Million cell updates/sec

Title: US-09-674-824-1

Perfect score: 2771

Sequence: 1 cgcactccactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*

1:	/SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2771	100.0	2805	21	AAZ24487
2	2541.8	91.7	2662	20	AAZ34651
3	2239	80.8	2239	19	AAV01527
4	1393	50.3	2383	19	AAT95785
5	1392.6	50.3	2491	21	AAZ50636
6	1391.4	50.2	2008	21	AAZ50643
7	1332	48.1	2990	19	AAV66832
8	1329.4	48.0	2533	15	AAQ45183
9	1327.6	47.9	2992	18	AAQ67285
10	1324	47.8	1752	19	AAV29756
11	1322.4	47.7	1749	19	AAV70960

C	12	1248.4	45.1	1620	19	AAV29759	Zea mays pEXS52 st
	13	1215.8	43.9	1328	21	AAZ50637	Corn soluble starch
	14	1015	36.6	1415	21	AAZ50642	Corn soluble starch
	15	648.4	23.4	1758	17	AAT32325	Soluble starch syn
	16	570.8	20.6	1984	21	AAZ46797	Arabidopsis thalia
	17	567.6	20.5	10336	20	AAZ34652	Wheat starch solub
	18	349	12.6	2380	19	AAV66834	Zea mays soluble s
	19	348	12.6	2478	18	AAT67287	Soluble starch syn
	20	299	10.8	5072	20	AAZ34653	Wheat starch solub
	21	233	8.4	2097	19	AAV29755	Zea mays soluble s
	22	233	8.4	2423	19	AAV70958	DNA encoding maize
C	23	232	8.4	1798	21	AAZ50647	Corn starch syntha
	24	232	8.4	2019	21	AAZ50646	Corn starch syntha
	25	232	8.4	2248	21	AAZ50651	Corn starch syntha
	26	208.6	7.5	2939	21	AAZ50647	Wheat starch synth
	27	208.6	7.5	2939	21	AAZ50647	Wheat starch synth
	28	195.8	7.1	2807	21	AAZ50647	Wheat starch synth
	29	195.8	7.1	2807	21	AAZ50647	Wheat starch synth
	30	193.2	7.0	2826	19	AAV01528	Wheat granule-boun
	31	190.8	6.9	2007	19	AAV70959	DNA encoding maize
	32	190.8	6.9	2007	19	AAV70959	Zea mays soluble s
	33	190.8	6.9	2085	18	AAZ67286	Soluble starch syn
	34	190.8	6.9	2085	19	AAZ66833	Zea mays soluble s
	35	186.8	6.7	2107	21	AAZ66833	Wheat starch synth
	36	186.8	6.7	2107	21	AAZ66833	Wheat starch synth
	37	172.4	6.2	9024	21	AAZ66833	Wheat SSII gene SE
C	38	155.8	5.6	11384	19	AAV52178	Streptococcus pneu
C	39	146.6	5.3	2571	19	AAZ96227	S. pneumoniae deri
	40	137.6	5.0	2115	21	AAZ32824	Arabidopsis thalia
	41	109.8	4.0	2542	19	AAV29753	Oryza sativa waxy
	42	104.4	3.8	1275	22	AAZ00044	Wheat cDNA encodin
C	43	103.4	3.7	2652	18	AAT75137	DNA encoding a 2.
C	44	102	3.7	396	22	AAZ94862	Human ovarian canc
C	45	100.2	3.6	612	22	AAH71471	Human cervical can

ALIGNMENTS

RESULT 1

AAZ24487

ID AAZ24487 standard; DNA; 2805 BP.

XX

AC AAZ24487;

XX

DT 18-FEB-2000 (first entry)

XX

DE Wheat soluble starch synthase DNA.

XX

KW Soluble; starch synthase; wheat; transgenic plant; starch production;

KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;

KW adhesive; textile; building material; soil stabilizer; wetting agent;

KW fertilizer; plant-protection; cosmetic; flocculant; ss.

XX

OS Triticum aestivum.

XX

FH Key Location/Qualifiers

FT CDS 314..2584

FT /*tag= a

FT /product= "soluble starch synthase"

XX

PN DE19820607-A1.

XX

PD 11-NOV-1999.

XX

PF 08-MAY-1998; 98DE-1020607.

XX

PR 08-MAY-1998; 98DE-1020607.

XX

PA (AGRE) HOECHST-SCHERING AGREVO GMBH.

XX

PI Loerz H, Luetticke S, Block M;

XX

WPI; 2000-024508/03.
P-PSDB: AAY50818.

P-PSDB; AAY50818.

New enzyme with starch synthase activity, useful for producing starch for foods and packaging materials -

for foods and packaging materials -

[illegible]

This invention describes a novel protein (I) with the activity of wheat starch synthase. Transgenic plants, specifically wheat, that contain (I) are used for production of starch, used particularly in foods, particularly baked and pastry goods and for making packaging materials or disposable items. Starch may also be used as starting materials for glucose or glucan components (e.g. for fermentation or further chemical conversion); in paper and pulp production, as adhesives, in textiles, in preparation of gypsum-based building materials, as soil stabilizer, as wetting agent etc. in fertilizer and plant-protection compositions, as binder (in pharmaceuticals, cosmetics, coal briquetting and casting sand), as flocculant in soil or coal slurries, as rubber and leather additives, and for production of synthetic polymers, e.g. polyurethane films. Transgenic plants with increased/decreased production of (I) produce starches with altered physical and/or chemical properties such as amylose/amylopectin ratios, degree of branching, mean chain length, phosphate content, gelatinization properties, gel- or film-forming properties, or starch grain size or structure. This sequence encodes the soluble starch synthase isolated from wheat (*Triticum aestivum* L. cv. Florida).

Sequence 2805 BP: 683 A: 703 C: 763 G: 656 T: 0 other;

Query Match	100.0%;	Score 2771;	DB 21;	Length 2805;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2771;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy	1	cgcacatcaactcgcttgcccacattcccaacttctctccccggcagcacacgaagtgcgc	60
Db	35	cgcacatcaactcgcttgcccacattcccaacttctctccccggcagcacacgaagtgcgc	94
Qy	61	accggtcatcacccataccttcgctcggccacacggcaaaccccccgatcgcgttttg	120
Db	95	accggtcatcacccataccttcgctcggccacacggcaaaccccccgatcgcgttttg	154
Qy	121	caggcagcgcaataaaacccccggggagcgcccccgcgacgacgacgacgacgtggg	180
Db	155	caggcagcgcaataaaacccccggggagcgcccccgcgacgacgacgacgtggg	214
Qy	181	agaagagaggttcggccccggccacgacagcggggcgatcacacgttcggtccgcaca	240
Db	215	agaagagaggttcggccccggccacgacagcggggcgatccacacgttcggtccgcaca	274
Qy	241	cctcctcgcctcctcctccttgctccgcgcgcacacacccatggcgcgacggcgctgcgc	300
Db	275	cctcctcgcctcctcctccttgctccgcgcgcacacacccatggcgcgacggcgctgcgc	334
Qy	301	gccgggtgcctgcgccccaagctccgccttgcgcgcgacatccggcgacggcgcccgggcg	360
Db	335	gccgggtgcctgcgccccaagctccgccttgcgcgcgacatccggcgacggcgcccgggcg	394
Qy	361	tccgccttcgcttcgcgcgcgcgccttcggcgcgccgccttcgcgcgggcgcttacgtccgcag	420
Db	395	tccgccttcgcttcgcgcgcgcgccttcggcgcgccgccttcgcgcgggcgcttacgtccgcag	454
Qy	421	ctcagcagggaggccccgc	480
Db	455	ctcagcagggaggccccgc	514
Qy	481	gtgcacaggttccttcgcag	540
Db	515	gtgcacaggttccttcgcag	574
Qy	541	ccgccccctgcgcgagccggcgctgggggaaactcgccccgaacctcgtctgaagggat	600

QY 1681 ggcatgatctcattaaatggccattccagagctcatgaggagagcgtgcaatttgc 1740
 |||||
 Db 1715 ggcattgatctcattaaatggccattccagagctcatgaggagagcgtgcaatttgc 1774
 |||||
 QY 1741 atgcttgatctggggtatcccaatttttgaagcgtgatgagatctaccagtcagttac 1800
 |||||
 Db 1775 atgcttgatctggggtatcccaatttttgaagcgtgatgagatctaccagtcagttac 1834
 |||||
 QY 1801 aagataaattccgtgagtggttgatttagttccagtttccacagaataactgca 1860
 |||||
 Db 1835 aagataaattccgtgagtggttgatttagtttccagtttccacagaataactgca 1894
 |||||
 QY 1861 ggttcgataatgttaattgcctatgagatttgaaccttgccgttcttaactgctatat 1920
 |||||
 Db 1895 ggttcgataatgttaattgcctatgagatttgaaccttgccgttcttaactgctatat 1954
 |||||
 QY 1921 gctatgaatatggtacagttcctgttagttcatggaactggggccctccgagacacagtc 1980
 |||||
 Db 1955 gctatgaatatggtacagttcctgttagttcatggaactggggccctccgagacacagtc 2014
 |||||
 QY 1981 gagacttcaaccttttgggtgcaaaagagagagaggttacaggttggtggttctcacgc 2040
 |||||
 Db 2015 gagacttcaaccttttgggtgcaaaagagagagaggttacaggttggtggttctcacgc 2074
 |||||
 QY 2041 ctacacgttgacaagatgtttgtggcatttgcgaaccgcgatgtcgcacattcaggaggcac 2100
 |||||
 Db 2075 ctacacgttgacaagatgtttgtggcatttgcgaaccgcgatgtcgcacattcaggaggcac 2134
 |||||
 QY 2101 aagccctcctggagggttcattgaagcagagcagatgacgaaagacattacgtgggacat 2160
 |||||
 Db 2135 aagccctcctggagggttcattgaagcagagcagatgacgaaagacattacgtgggacat 2194
 |||||
 QY 2161 gcccagagcagatcagagcagattcagtggtggtccttgtagcacaacctacgtcatgt 2220
 |||||
 Db 2195 gcccagagcagatcagagcagattcagtggtggtccttgtagcacaacctacgtcatgt 2254
 |||||
 QY 2221 agacgggactggagggttccaaagtcagtcctcctcagctctgaagacatcctctca 2280
 |||||
 Db 2255 agacgggactggagggttccaaagtcagtcctcctcagctctgaagacatcctctca 2314
 |||||
 QY 2281 tccttcgggcccgaagataccctctacattgctgttctctcagtagagtcg 2340
 |||||
 Db 2315 tccttcgggcccgaagataccctctacattgctgttctctcagtagagtcg 2374
 |||||
 QY 2341 caatgcgctgtgttgggttcgcggttcgagacatatgaacggtgtgtgtgtgcg 2400
 |||||
 Db 2375 caatgcgctgtgttgggttcgcggttcgagacatatgaacggtgtgtgtgtgcg 2434
 |||||
 QY 2401 gcggtgacagcttcgggtgacacagcttacagtttggggaataaggaaggtatgtct 2460
 |||||
 Db 2435 gcggtgacagcttcgggtgacacagcttacagtttggggaataaggaaggtatgtct 2494
 |||||
 QY 2461 gcaggtatggttaacagcaagcaccactcagatggcagcctctctctcgtgtttacagct 2520
 |||||
 Db 2495 gcaggtatggttaacagcaagcaccactcagatggcagcctctctctcgtgtttacagct 2554
 |||||
 QY 2521 gaaatcagaacacactggtgactcttttagccttagttagtgaagtttggcttctct 2580
 |||||
 Db 2555 gaaatcagaacacactggtgactcttttagccttagttagtgaagtttggcttctct 2614
 |||||
 QY 2581 gttatgttcttcttcttagctgacaaatatttgacctgttggagaattttatctttg 2640
 |||||
 Db 2615 gttatgttcttcttcttagctgacaaatatttgacctgttggagaattttatctttg 2674
 |||||
 QY 2641 ctgctgttttttttaatacaaaagaggggtttctcccgatatttcatttaaaaaa 2700
 |||||
 Db 2675 ctgctgttttttttaatacaaaagaggggtttctcccgatatttcatttaaaaaa 2734
 |||||
 QY 2701 aaaaaa 2760
 |||||
 Db 2735 aaaaaa 2794
 |||||

QY 2761 aaaaaaaaaa 2771
 |||||
 Db 2795 aaaaaaaaaa 2805
 |||||
 RESULT 2
 AAX34651
 ID AAX34651 standard; cDNA; 2662 BP.
 XX
 AC AAX34651;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 cDNA sequence of wheat starch soluble synthase I (SSS I).
 XX
 Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;
 KW starch branching enzyme; starch soluble synthase; debranching enzyme;
 KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
 KW grain softness protein I; bacterial isoamylase; glycogen synthase;
 KW WSE I-D4 gene; ss.
 XX
 OS Triticum tauschii.
 XX
 PN WO914314-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 11-SEP-1998; 98WO-AU00743.
 XX
 PR 20-MAR-1998; 98AU-0002509.
 PR 12-SEP-1997; 97AU-0009108.
 XX
 (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Li Z, Morell M, Rahman S;
 XX
 DR WPI; 1999-229525/19.
 XX
 PT New isolated cereal plant enzyme genes used for, e.g. expression of
 PT antisense sequences of granule bound synthase
 XX
 PS Claim 10; Page 81-83; 171pp; English.
 XX
 CC The invention relates to a novel enzyme of starch biosynthetic pathway
 CC in a cereal plant, where the enzyme is selected from starch branching
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
 CC SBE I of rice or maize. The methods and products can be used for
 CC targeting expression specifically to the endosperm of the seeds of cereal
 CC plants such as wheat or barley. They can be used for the expression of
 CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
 CC can be used for modifying the characteristics of starch produced by a
 CC plant. The present sequence represents the wheat SSS I cDNA sequence.
 XX
 SQ Sequence 2662 BP; 592 A; 677 C; 754 G; 639 T; 0 other;

Query Match 91.7%; Score 2541.8; DB 20; Length 2662;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 2597; Conservative 0; Mismatches 32; Indels 6; Gaps 3;
 QY 25 ctccactcttctctcccgccacacccgagtcgagcggcgcgcacccatcacctcg 84
 |||||
 Db 2 ctccactcttctctcccgccacacccgagtcgagcggcgcgcacccatcacctcg 61
 |||||
 QY 85 gctcggccaccggcaaaccccccgatccgcttttgcaggcagcagcactaaacccggg 144
 |||||
 Db 62 gctcggccaccggcaaaccccccgatccgcttttgcaggcagcagcactaaacccggg 121
 |||||

for production of transgenic plants, particularly starch-producing plants, specifically wheat. Use of the isolated nucleic acids, or of antisense sequences, allows starch metabolism to be regulated in transgenic plants. Overexpression may result in improved crop yield, while modification of starch in planta may eliminate the need for subsequent chemical/physical modification. Plants with altered levels of the various isoforms of starch synthase will produce starch of different chain length, amylose/amylopectin ratio, degree of branching, phosphate content, gelatinisation behaviour, granule size and shape, viscosity etc. The starch produced by such plants is useful particularly in foods or to produce packaging materials or disposable goods, as well as in any other known use of starch.

Sequence 2239 BP; 611 A; 448 C; 590 G; 590 T; 0 other;

Query Match 80.8%; Score 2239; DB 19; Length 2239;
Best Local Similarity 100.0%; Pred. No. 1.7e-311;
Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 533 cgagcagccgcccctgcgagccgcccgcggggaactgcgcgcgacccctcctgctcg 592
Db 1 cgagcagccgcccctgcgagccgcccgcggggaactgcgcgcgacccctcctgctcg 60
Qy 593 aagggtgctgaggtatccatcgacagcataattgtggtgcgaagtgcagcagattctg 652
Db 61 aagggtgctgaggtatccatcgacagcataattgtggtgcgaagtgcagcagattctg 120
Qy 653 agatcatggtgcgaatgagcaacctcaagctaaagtacacgtacgtctgttctgga 712
Db 121 agatcatggtgcgaatgagcaacctcaagctaaagtacacgtacgtctgttctgga 180
Qy 713 ctggtgaagctgctcctctatgcgaagtcagggggtggtggagatgtttgtggtctg 772
Db 181 ctggtgaagctgctcctctatgcgaagtcagggggtggtggagatgtttgtggtctg 240
Qy 773 caattgctctgctgctggtgcacccagtgatggtgtgtaatgccagatactaaatg 832
Db 241 caattgctctgctgctggtgcacccagtgatggtgtgtaatgccagatactaaatg 300
Qy 833 ggtcctctgataaaactatgcgaagtcataacacgcgaagcacataagattccat 892
Db 301 ggtcctctgataaaactatgcgaagtcataacacgcgaagcacataagattccat 360
Qy 893 gctttggggatcacatgaagtgaacttttctcatgagtagatagagacacgctgattgg 952
Db 361 gctttggggatcacatgaagtgaacttttctcatgagtagatagagacacgctgattgg 420
Qy 953 tgtttgctgatacctgctatcatcacagaccaggaagtttatatgagataattttggtg 1012
Db 421 tgtttgctgatacctgctatcatcacagaccaggaagtttatatgagataattttggtg 480
Qy 1013 cttttggtgataatcagtttcagatacacactcctttgctatgctgcatgcgaggccccc 1072
Db 481 cttttggtgataatcagtttcagatacacactcctttgctatgctgcatgcgaggccccc 540
Qy 1073 taatccttgaattggaggagatatttatggacagaaattgctgtttgtggaacgatt 1132
Db 541 taatccttgaattggaggagatatttatggacagaaattgctgtttgtggaacgatt 600
Qy 1133 ggcagtcgagccttggcagctcctctctgctgcaaaatagacacacgctgtttaca 1192
Db 601 ggcagtcgagccttggcagctcctctctgctgcaaaatagacacacgctgtttaca 660
Qy 1193 gagattccccgagcaccccttcttatataataatttagcacatcaggtgtggaacctgcaa 1252
Db 661 gagattccccgagcaccccttcttatataataatttagcacatcaggtgtggaacctgcaa 720
Qy 1253 gtacatactctgattcctggtggtggtggtggtggtggtggtggtggtggtggtggtg 1312
Db 721 gtacatactctgattcctggtggtggtggtggtggtggtggtggtggtggtggtggtg 780

Db 2282 cccctgtacattgctgtcctgctacagtagatcgcaatgctgctgctgctgctgctgctg 2340
Qy 2364 ccccggttcgagacaatgagcggctgtgctgctgctgctgctgctgctgctgctgctgctg 2423
Db 2341 cgcgggttcgagtagatgagcggctgtgctgctgctgctgctgctgctgctgctgctgctg 2400
Qy 2424 acagttacagttttggggaataaggaaggatgctgctgctgctgctgctgctgctgctgctg 2483
Db 2401 acagttacagttttggggaataaggaaggatgctgctgctgctgctgctgctgctgctgctg 2460
Qy 2484 ccactcagatggcagcctcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2543
Db 2461 ccactcagatggcagcctcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2520
Qy 2544 tctttagccttagtgaattggaattgctgctgctgctgctgctgctgctgctgctgctgctg 2603
Db 2521 tctttagccttagtgaattggaattgctgctgctgctgctgctgctgctgctgctgctgctg 2580
Qy 2604 tgacaaattatgacctgttggaagaat---tttatcttctgctgctgctgctgctgctgctg 2654
Db 2581 tgacaaattatgacctgttggaagaat---tttatcttctgctgctgctgctgctgctgctg 2635

RESULT 3

AAV01527
ID AAV01527 standard; cDNA to mRNA; 2239 BP.

XX AC AAV01527;
XX DT 21-MAY-1998 (first entry)

XX DE Wheat soluble starch synthase partial-cDNA sequence.
XX KW Starch synthase; wheat; transgenic plant; ss.
XX OS Triticum aestivum L. cv. Florida.

XX FH Key Location/Qualifiers
XX FT CDS 3..2018
XX FT /*tag= a

XX PN W09745545-A1.
XX PD 04-DEC-1997.

XX PF 28-MAY-1997; 97WO-EP02793.
XX PR 11-SEP-1996; 96DE-1036917.

XX PR 29-MAY-1996; 96DE-1021588.
XX PA (AGRE) HOECHST-SCHERING AGREVO GMBH.

XX PI Block M, Loerz H, Luetticke S, Frohberg C, Kossmann J;
XX PI Walter L;

XX DR WPI: 1998-032652/03.
XX DR P-PSDB; AAW23937.

XX Nucleic acid encoding starch synthase enzymes from wheat - for
PT transgenic plants that produce modified forms of starch, useful e.g.
PT in foods, or for production of packaging materials and disposable
PT goods
XX Claim 1; Page 47-51; 71pp; English.

XX This near full-length cDNA clone, designated TaSSS, codes for a
CC soluble starch synthase (see AAW23837) of summer wheat (cv. Florida).
CC It was isolated from a phage cDNA library of 21-day-old wheat
CC caryopses by screening with a PCR fragment derived from rice soluble
CC starch synthase (see also AAV01529-30). A second clone (see AAV01528),
CC coding for wheat granule-bound starch synthase (see AAW23938) is also
CC claimed. These isolated nucleic acids can be inserted into vectors

```
QY 1313 cagaatggcaggagacatgcctctgacaaagggtgagcagtaactttttgaaaggag 1372
    |||||
Db 781 cagaatggcaggagacatgcctctgacaaagggtgagcagtaactttttgaaaggag 840
    |||||
QY 1373 cagttgacagcagatcgagttgtgacccgtcagtcaggttattcatgggaggtcacaa 1432
    |||||
Db 841 cagttgacagcagatcgagttgtgacccgtcagtcaggttattcatgggaggtcacaa 900
    |||||
QY 1433 ctgctgaaggtgacagggcctcaa tgaagctcttaagctccccgaaaaagtcatigaatg 1492
    |||||
Db 901 ctgctgaaggtgacagggcctcaa tgaagctcttaagctccccgaaaaagtcatigaatg 960
    |||||
QY 1493 gaattgtaaatgaattgacattaaatgattggaacccccacacagacaagtctccctc 1552
    |||||
Db 961 gaattgtaaatggaattgacattaaatgattggaacccccacacagacaagtctccctc 1020
    |||||
QY 1553 atcattattctgtcagtgacccctctctggaaggcccaaatgtaaaagctaaattgcaagaag 1612
    |||||
Db 1021 atcattattctgtcagtgacccctctctggaaggcccaaatgtaaaagctaaattgcaagaag 1080
    |||||
QY 1613 agttgggttaacctgtaagggaggtgttccctctgattggtttattggaagactggatt 1672
    |||||
Db 1081 agttgggttaacctgtaagggaggtgttccctctgattggtttattggaagactggatt 1140
    |||||
QY 1673 accagaaaggcattgatctcattaaatggaatggccattccagagctcatgaggagagacgtgc 1732
    |||||
Db 1141 accagaaaggcattgatctcattaaatggaatggccattccagagctcatgaggagagacgtgc 1200
    |||||
QY 1733 aatttgatcgttgatctggtggggtcccaatttttgaaggctggtgagatcaccaggt 1792
    |||||
Db 1201 aatttgatcgttgatctggtggggtcccaatttttgaaggctggtgagatcaccaggt 1260
    |||||
QY 1793 cgagttacaagataaaatccgtgagatgggttgatttagttccaggtttccacagaa 1852
    |||||
Db 1261 cgagttacaagataaaatccgtgagatgggttgatttagttccaggtttccacagaa 1320
    |||||
QY 1853 taactgaggttgagatatattgtaatgccatgagatttgaaacctggcgttccaaac 1912
    |||||
Db 1321 taactgaggttgagatatattgtaatgccatgagatttgaaacctggcgttccaaac 1380
    |||||
QY 1913 agctatatgctacaatggttacagttcctctgattgatttcagaaactggggcctccag 1972
    |||||
Db 1381 agctatatgctacaatggttacagttcctctgattgatttcagaaactggggcctccag 1440
    |||||
QY 1973 acacagtcgagacctcaacccctttggtgcaaaaggagagaggtacaggtggcgt 2032
    |||||
Db 1441 acacagtcgagacctcaacccctttggtgcaaaaggagagaggtacaggtggcgt 1500
    |||||
QY 2033 tctcaccgtcaacccgtcctgaggaaggtctgaggaaggtggaacccgcgagtgacattca 2092
    |||||
Db 1501 tctcaccgtcaacccgtcctgaggaaggtctgaggaaggtggaacccgcgagtgacattca 1560
    |||||
QY 2093 gggagcacagccgtcctgaggaaggtctgaggaaggtggaacccgcgagtgacattca 2152
    |||||
Db 1561 gggagcacagccgtcctgaggaaggtctgaggaaggtggaacccgcgagtgacattca 1620
    |||||
QY 2153 gggacacatcccagcagcagatcagagatcttcagtgaggcctctgagcaacccta 2212
    |||||
Db 1621 gggacacatcccagcagcagatcagagatcttcagtgaggcctctgagcaacccta 1680
    |||||
QY 2213 cgtcatgtagaggggactgggaggttccagtcagagtcctctcagctctgaagacat 2272
    |||||
Db 1681 cgtcatgtagaggggactgggaggttccagtcagagtcctctcagctctgaagacat 1740
    |||||
QY 2273 cctcttcaccttcggcccggaagataccctgtacattgctgttctcgtctacag 2332
    |||||
Db 1741 cctcttcaccttcggcccggaagataccctgtacattgctgttctcgtctacag 1800
    |||||
QY 2333 tagagtcgaaatgcgcctgctgtctgttctgctgcgggttcagaaacatagacggtgtg 2392
    |||||
Db 1801 tagagtcgaaatgcgcctgctgtctgttctgctgcgggttcagaaacatagacggtgtg 1860
    |||||
QY 2393 ctgctgcggcgtgacagcttcgggtgagcagacagttacagtttttggggaataaggag 2452
    |||||
```

RESULT 4

AAT95785

ID AAT95785 standard; cDNA to mRNA; 2383 BP.

XX AC AAT95785;

XX DT 22-MAY-1998 (first entry)

XX DE Maize starch synthase type I cDNA.

XX KW Maize; starch synthase type I; starch; ds.

XX OS Zea mays.

XX FH Key Location/Qualifiers

XX FT CDS 2..1951

XX FT /*tag= a

XX FT /product= starch_synthase_type_I

XX PN DE19619918-A1.

XX PD 20-NOV-1997.

XX PF 17-MAY-1996; 96DE-1019918.

XX PR 17-MAY-1996; 96DE-1019918.

XX PA (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.

XX FI Froberg C, Kossmann J;

XX DR WPI; 1998-000821/01.

XX DR P-PSDB; AAW38218.

XX PT DNA encoding maize starch synthase type I protein - for producing transgenic plants

XX PS Claim 1; Pages 16-20; 23pp; German.

XX CC The present sequence encodes maize starch synthase type I, useful in the production of starch. Starch can be used in various conventional starch applications, e.g. starch hydrolysate products, CC foods, papermaking, adhesives, textiles, building materials, soil CC stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal CC briquettes, ore and coal slurries, foundry casting, rubber, leather

Db 1494 aaaggcatgatcatgaataaattgcattccagatctcatcgcgacaaattcaatc 1553
QY 1738 gtcatcttgatctggagatccaatttttgaaggctgagatgagatctaccgagtcagt 1797
Db 1554 gtcatcttgatctggatcgaccaggttttgaaggatgagatgagatccacagaatcagg 1613
QY 1798 tacagagataaattccgttgatgggttgatttagttgctccagttccacagaataact 1857
Db 1614 tacagagataaattccgttgatgggttgatttagttgctccagttccacagaataact 1673
QY 1858 gaaggttcgatatattgttaatgcacatgcagagatttgaacctgctgttttaatacagcta 1917
Db 1674 gcaggttcgatatattgttgatccatccagatttcgaacctgtgacctcaatcagcta 1733
QY 1918 tatgctatcaatatgtatcagttcctgttagttcatggaactggggcctccgagacaca 1977
Db 1734 tatgctatgcaatatgtatcagttcctgttgcttcatggaactggggcctccgagatata 1793
QY 1978 gtgagacattcaacctttttgtgcataaaggagagagaggttacaggtgggcttctca 2037
Db 1794 gtgagaattttaacctgttctgagaaaggagagagaggttacaggtgggcttctcg 1853
QY 2038 cegtaacc-gtggacaaagattgttggtcattgcgaacccgagatgcacattcaggga 2096
Db 1854 ccaataaccattgaaataatgctgtggcattgcgattggcaatttcgacatacaggga 1913
QY 2097 gcacaagccgtcctggaggggtcctatgaagcgagcagatgacgaagaccatcgtggga 2156
Db 1914 acacaagtcctctggagggttctaataagcgagagcagatgctcaagcgactttaca 1973
QY 2157 ccattgcc-ccgagcagatcagacagatcttcagttgggcttctggtgacacacccatcgt 2215
Db 1974 ccaatgcgctcacagatgaaacagatcttcgaatggccttcattgatcaaccatagt 2033
QY 2216 catgtagacgg 2226
Db 2034 catgtaaatgg 2044

RESULT 9
AAT67285
ID AAT67285 standard; cDNA; 2992 BP.
XX
AC AAT67285;
XX
DT 11-SEP-1997 (first entry)
XX
DE Soluble starch synthase cDNA clone SSS10.52.
XX
KW Soluble starch synthase; SSS10.52; transgenic plant; cereal; maize;
KW ss.
XX
OS Zea mays inbred line W64A.
XX
PN W09720936-Al.
XX
PD 12-JUN-1997.
XX
PF 04-DEC-1996; 96WO-GB02990.
XX
PR 06-DEC-1995; 95GB-0024938.
XX
PA (ZENE) ZENECA LTD.
XX
PI Keeling PL, Knight ME;
XX
DR WPI: 1997-319782/29.
XX
PT cDNA encoding soluble starch synthase - used to produce transgenic
PT plants with increased capacity for producing and storing starch
XX
PS Claim 5; Page 17-20; 44pp; English.

XX
CC cDNA clones SSS10.52, SSS6.31 and SSS56 (AAT67285-87) code for maize
CC soluble starch synthase (SSS). They were isolated from a maize
CC inbred line W64A library by screening with a probe (AAT67288) based
CC on pea SSS, rice SSS and maize GBSS. The cDNAs can be used to
CC produce plants with an increased capacity for producing starch, or
CC a capacity to produce starch with an altered fine structure. They
CC can also be used to isolate the corresponding genomic sequences
CC from crop plants, to determine the contribution of the SSS gene to
CC the net regulation of starch biosynthesis, and to modify the levels
CC of starch produced by the plant. Transgenic plants (esp. maize)
CC can be used to produce hybrid plants which have higher rates of
CC starch synthesis at temperatures above the normal optimum.
XX
SQ Sequence 2992 BP; 758 A; 655 C; 801 G; 776 T; 2 other;

Query Match 47.9%; Score 1327.6; DB 18; Length 2992;
Best Local Similarity 82.0%; Pred. No. 2e-181;
Matches 1606; Conservative 0; Mismatches 335; Indels 18; Gaps 6;

QY 247 cgcctctctccctgtcccgcgccacacccatggcgcgagcggtcgcgcgcg 306
Db 563 ccccccgcgcgcgtctgtctctcctcctcgcgaatggcgagcgcctcgccgtg 622
QY 307 tgctcgccccagcgtccgcgcgtcgccgacccgacgagcgccggcggtccgcc 366
Db 623 ggcgcgctgctctctcctcgcggcgctgcccgcgcctcg-----cgac 676
QY 367 tgctgctcgcgcgcgctcccgcttgcgcgggcgctacgtcgcgagctcagc 426
Db 677 cgggcgcccgcgaggtccagcgtgctgcgcgcgctgctgcggagctgagc 736
QY 427 agggagggcccgcgcgcccgcgagcagcagcagcagcagcagcagcagcagc 486
Db 737 agggagggcccgcgcgcccgcgctgctgcgcgcgcgcgcgcgcgcgcgcgcgc 796
QY 487 ggctctcgc 546
Db 797 ggctctcg---cgcgcgcccgcgagccacggtgagccgacgagcgcgccgc 853
QY 547 ctgcccgcgcgcgcgtgggggaactcgcgccgcgcgcgcgcgcgcgcgcgcgcgc 606
Db 854 gtgcccgcgcgcgcgtgggggacctgctcga-----acctgaaggattgctga 907
QY 607 gattccatgcagacataattgtgctgcaagtgcagagcagagattctgagatcgag 666
Db 908 ggtccatgcataacacagtagttgtggcaagtgcagagattctgagattggttgg 967
QY 667 aatgcaaacctcaagctaaagtacacgtagcatcggtttgtgactggtgagctgct 726
Db 968 aaggagcagctcagctaaagttaacacaaagcatgctcttgaaccggaagcttct 1027
QY 727 cctt-atgcaaaatcaggggggttgggagatgttggctgcttaccacattgctctgc 785
Db 1028 ccttaatcgaaagctgggggtctaggagatgttggcttattgccaattgctctgc 1087
QY 786 tgctcggttcaccgagtgatgttgaatgccaagatacttaaatgggtccctcgataa 845
Db 1088 tgctcggttcaccggtgtgattgtaatgcccagacatttaaatggtaccctccgataa 1147
QY 846 aaactatgcaaggcattatcacactgcgaagcacataagattccatgcttgggggac 905
Db 1148 gaattatgcaaatgcaatttactcagaaaaaacacattcgagattccatttggcggtga 1207
QY 906 acatgaagtgcaccttttctaatgagtatagagacaacgtcgattgggtgtttgcagta 965
Db 1208 acatgaagttaaccttctccatgagtatagatgagttgactgactgggtgtttgtgata 1267
QY 966 tccgtatatcacagaccagggaagtattatgagataaatttgggtgcttttgggtgataa 1025
Db 1268 tccctatatcacagacctggaaatttatatggagataaatttgggtgcttttgggtgataa 1327


```
RESULT 12
AAV29759
ID AAV29759 standard; DNA; 1620 BP.
XX
AC AAV29759;
XX
DT 11-SEP-1998 (first entry)
XX
DE Zea mays pEXS52 starch synthase gene.
XX
KW SER; starch-encapsulating region; fusion vector;
KW starch synthase; glucosyl transferase; pEXS52; ss.
XX
OS Zea mays.
XX
FH Key
FT Location/Qualifiers
FT 1..2007
FT /*tag= a
FT /product= starch synthase
XX
PN W09814601-A1.
XX
PD 09-APR-1998.
XX
PF 30-SEP-1997; 97WO-US17555.
XX
PR 30-SEP-1996; 96US-0026855.
XX
PA (EXSE-) EXSEED GENETICS LLC.
XX
PI Guan H, Keeling P;
XX
DR WPI: 1998-240100/21.
DR P-PSDB; AAW56491.
XX
XX Hybrid polypeptide comprising starch-encapsulating region and
PT protein - useful for, e.g. producing protein(s) resistant to
PT degradation by stomach acids
XX
PS Example 5; Page 51-53; 156pp; English.
XX
CC The sequence is that of the starch synthase coding region from
CC pEXS52. It can be used in the production of a hybrid polypeptide
CC comprising a starch-encapsulating region (SER) fused
CC to a payload protein. The hybrid polypeptide can be used to make
CC modified starches comprising the payload protein, selected from,
CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
CC immunoglobulins, etc. The modified starch can also be used
CC to provide grain feeds enriched in amino acids. By encapsulating
CC the payload protein in starch, it is more resistant to
CC degradation by stomach acids.
XX
SQ Sequence 1620 BP; 426 A; 313 C; 422 G; 459 T; 0 other;
```

```
Query Match 45.1%; Score 1248.4; DB 19; Length 1620;
Best Local Similarity 86.2%; Pred No. 3.9e-170;
Matches 1392; Conservative 0; Mismatches 221; Indels 1; Gaps 1;
```

```
QY 553 gacgcgcgcgtggggaactcgcgcgcacccctcctgctcgaaggaggtgctgaggtacc 612
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
4 gtcgcgcgcgtgacgagggagacccctcgtcgaacctgaaggaggtgctgaggtacc 63
QY 613 atcgacagcataatgtggtcgcagtgagcaggattctgagatcatgatgcggaatgag 672
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
64 atcgatacacagatgtgtggaagtgacgaagattctgagattgtggtggaaggag 123
QY 673 caacctcaagctaaagtacacgtagcatcgtgtgttgactggtgaactggtgaactccttat 732
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
124 caagctcagctaaagtacacaagacgtatccttctgtgaaccgcgcgaagcttctctat 183
```

```
QY 733 gcaaaagtccaggggggttgaggagatgtttgtggttcgttaccacaaatgtctctgtctcgt 792
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
184 gcaaaagtctgggggtctaggagatgtttgtggttcatttccacagttcgtctctgtctcgt 243
QY 793 ggtcaccgcgtgagtggttgtaatgccaaatacttaaatgggttcctctgtataaaactat 852
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
244 ggtcaccgtgtagtggttgtaatgccagataatttaaatgtgtaccctccgataagaattat 303
QY 853 gcaaaaggcattatcacactcgcgaagcacattaaagattccattccttcttgggggacacatgaa 912
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
304 gcaaatgcattttacacagaaaacacattcggattccattcgtcttggcgtgacatgaa 363
QY 913 gtgacctttttcatgagtatagacacacgtcgtgattgggtgttgcgtcatccgtca 972
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
364 gttaccttctccatgagtatagagattcagttgactgggtgttgttgatcatccctca 423
QY 973 tatcacagaccagggaagtattatatggagataattttgggtccttttgggtgataatcagttc 1032
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
424 tatcacagacctggaaatttatatggagataaagtgttggtccttgggtgataatcagttc 483
QY 1033 agatacacactccttgcctatgctgcagcgagcccccactaactccttgaattgggagga 1092
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
484 agatacacactccttgcctatgctgcagcgctccttgccttgaattgggagga 543
QY 1093 tatattatggacagaattgcattgttgtgaacgattggcatgccagccttctgtgcca 1152
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
544 tatattatggacagaattgcattgttgtgcaatgattggcatgccagctcagtgcca 603
QY 1153 gtcccttctgctgcaaaatatagaccatcagcgtgtttacagagattcccgcgacacccctt 1212
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
604 gtcccttctgctgcaaaatatagaccatcagcgtgtttataaaagactcccgcgacattctt 663
QY 1213 gttatacataatttagcacatcagggtgtggagcctgcgaagtacatatcctgtatcgtgga 1272
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
664 gtaatacataatttagcacatcagggtgtgagcctgcgaagcacaatatcctgcactcgtgg 723
QY 1273 ttgcctcctcgaatggtatggagccttttagaattgggtatttccagaattggcgaaggagcat 1332
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
724 ttgcacactgaatggtatggagcctcctggagtggtgttattccctgaaatgggcgagggagcat 783
QY 1333 gcccttgacaaagggtgagcagtttaacttttgaaggagcagttgtgcacagcagatcgg 1392
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
784 gcccttgacaaagggtgagcagtttaacttttgaaggagtgagttgtgcacagcagatcga 843
QY 1393 attgtgacgcgtcagtcaggttattcatgggaggtcacaaactcgtgaaagtgacaggggc 1452
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
844 atcgtgactgcagtaagggttattcgtgggaggtcacaaactcgtgaaagtgacaggggc 903
QY 1453 ctcaatgagcctttaaagctcccgaagaaagtgtattgaatgggaattgaaatggaattgac 1512
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
904 ctcaatgagcctttaaagctcccgaagaaagtgtattgaacgggaattgaaatggaattgac 963
QY 1513 attaatgattggaacccccaccacagacaagtgtcctccatcatctattctgtcgatgac 1572
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
964 attaatgattggaacccccaccacagacaagtgtcctccatcatctattctgtcgatgac 1023
QY 1573 ctctctgaaaggcccaaatgtaagctgaattgcagaagaggtgggtttaccctgtaagg 1632
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1024 ctctctgaaaggcccaaatgtaagctgaattgcagaagaggtgggtttaccctataagg 1083
QY 1633 gaggatgttccctcgtgattggtttattggaagactggaatccacagaagagcattgattctc 1692
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1084 cctgattgtcctcgtgattggtttattggaaggttggaattatcagaagagcattgattctc 1143
QY 1693 attaaatggccattccagagctcatgagggagacgtgcaattttgtcatgcttggatct 1752
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1144 attcaacttatcataccagatctcatcgcgggaagatgttcaatttgcattgctggtgact 1203
QY 1753 gggggtccaatttttgaagcctggatgagatctaccagagtcgagttcacagagataattc 1812
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1204 ggtgacccagagctggaagattggatgagatctacagatctcatcttcagagataaattc 1263
QY 1813 cgtggatgggttggattttagtgttccagtttccacagataactgcaggttgcgataata 1872
```


Search completed: March 28, 2002, 15:01:12
Job time: 7416 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen. Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2002, 12:13:06 ; Search time 104.55 Seconds
(without alignments)
6002.588 Million cell updates/sec

Title: US-09-674-824-1
Perfect score: 2771
Sequence: 1 cgcactccactcgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	80.8	2239	4	US-09-196-390-1
2	1393	50.3	2383	4	US-09-192-909-1
3	1332	48.1	2990	1	US-08-572-951-1
4	1325.6	47.8	1752	3	US-08-941-445A-12
5	1248.4	45.1	1620	3	US-08-941-445A-20
6	752.8	27.2	2360	3	US-08-836-567-9
7	648.4	23.4	1758	3	US-08-836-567-3
8	349	12.6	2380	1	US-08-572-951-3
9	287	10.4	2793	3	US-08-836-567-7
10	279	10.1	1926	3	US-08-836-567-5
11	233	8.4	2097	3	US-08-941-445A-10
12	193.2	7.0	2825	4	US-09-196-390-5
13	190.8	6.9	2007	3	US-08-941-445A-8
14	190.8	6.9	2085	1	US-08-572-951-2
15	109.8	4.0	2542	3	US-08-941-445A-6
16	103.4	3.7	2652	4	US-09-115-704-1
17	97.8	3.5	2671	6	5168051-9
18	96.6	3.5	1641	1	US-08-300-903A-8
19	95.6	3.5	1098	3	US-09-248-335-35
20	94.8	3.4	1117	4	US-09-247-373B-33
21	94	3.4	1582	3	US-08-545-196B-10
22	94	3.4	1582	3	US-08-545-196B-12
23	94	3.4	2852	3	US-09-027-137-2
24	92.4	3.3	1733	3	US-09-073-569-1
25	92	3.3	1464	1	US-07-735-065-1
26	92	3.3	1464	1	US-08-469-202-11
27	92	3.3	1464	2	US-08-484-434C-11

28	91.2	3.3	1813	4	US-09-071-224-3	Sequence 3, Appli
29	90.8	3.3	1882	4	US-09-370-253-1	Sequence 1, Appli
30	90	3.2	1198	3	US-09-248-335-27	Sequence 27, Appl
31	90	3.2	1474	4	US-08-821-994-64	Sequence 64, Appl
32	89.8	3.2	2303	3	US-08-836-567-1	Sequence 1, Appli
33	89.8	3.2	4168	3	US-08-836-567-11	Sequence 11, Appl
34	88.8	3.2	140	1	US-08-628-417-5	Sequence 5, Appli
35	88.8	3.2	240	1	US-08-628-417-6	Sequence 6, Appli
36	88.6	3.2	2447	2	US-09-014-969-14	Sequence 14, Appl
37	88.4	3.2	630	1	US-08-185-414E-1	Sequence 1, Appli
38	88.4	3.2	1817	1	US-08-473-981A-5	Sequence 5, Appli
39	88.4	3.2	1817	2	US-08-474-087-5	Sequence 5, Appli
40	88.4	3.2	2082	2	US-08-785-310A-2	Sequence 2, Appli
41	88	3.2	144	1	US-08-702-344-26	Sequence 26, Appl
42	88	3.2	3410	4	US-09-020-956-110	Sequence 110, App
43	88	3.2	3410	4	US-09-030-607-110	Sequence 110, App
44	87.6	3.2	1066	1	US-08-157-101A-4	Sequence 4, Appli
45	87.2	3.1	1602	1	US-08-530-950-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-196-390-1
; Sequence 1, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196.390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2239 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum L.
; STRAIN: cv. Florida
; HAPLOTYPE: ca. 21 d Caryopses
; IMMEDIATE SOURCE:
; LIBRARY: cDNA library in pBluescript sk (-)
; CLONE: TaSSS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..2017
; US-09-196-390-1

```

Query Match 80.8%; Score 2239; DB 4; Length 2239;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 533 cgacgacgccccctgcgcgacgacgccccgggggaactcgcgccccgacctcctgctcg 592
D 1 CGACGACGCGCCCTGCCGACGCGCGCTGGGGAACTCGCGCGCGACCTCCTGCTCG 60
QY 593 aaggattgctgaggtatccatcgacagcataatgtggctgcaagtgcagcagattctg 652
D 61 AAGGGATTGCTGAGGATTCCATCGACAGCATATTTGGTGGCTGCAAGTGACGAGATTCTG 120
QY 653 agatcaygatgogaaatgacaaacctcaagctaaagtatacagtagcatcgtgttgtga 712
D 121 AGATCATGGATGGAATGAGCAACTCAAGCTAAAGTTACACGTAGCATCGTGTGTGA 180
QY 713 ctggtgaagctgctcctatgcaaaagcagggggtggagattgttggcttcgttac 772
D 181 CTGGTGAAGCTGCTCTTATGCAAAAGTCAGGGGGTGGGAGATGTTTGTGTTGCTTAC 240
QY 773 caattgctctgctgctggttcacgagtgatggttgaatgccaaagatacttaagt 832
D 241 CAATTGCTCTTGTGCTGCTGGTCACCGAGTGATGGTTGTAATGCCAAGATACTTAATG 300
QY 833 ggtcctctgataaaactatgcaaaagcattatatacctgcgaagcacataagattccat 892
D 301 GGTCTCTGATAAAACTATGCAAAAGCATATACACTGCGAAGCACATTAAAGATTCCAT 360
QY 893 gcttgggggacacatgaagtgaccttttctatgagtagagacaacgctcgattggg 952
D 361 GCTTTGGGGATCACATGAAGTGACCTTTTTCATGAGTATAGAGACAACGTCGATTGGG 420
QY 953 tgttgtcgatcatccgtctatatacacagaccaggaagtattatatggagataatttgg 1012
D 421 TGTTTGTCGATCATCCGTCATATCACAGACCAGGAAGTTTATATGAGAGATAAATTTGGTG 480
QY 1013 ctttgggtgataaacagttccagatacacacctccttctatgctgcagcagggccccac 1072
D 481 CTTTGGTGAATAACGTTTCAGATACACACTCTCTTGTCTATGCTGCAATGCGAGGCCCCAC 540
QY 1073 taaccttgattggaggatatttatggacagaattgcatgtttgttgtgaacgatt 1132
D 541 TAATCCTTGAAATGGGAGGATATATTTATGGACAGAAATTCATGTTTGTGTGAACGATT 600
QY 1133 ggcatgccagcttgtccagctccttctgctgcaaaatagaccatacaggtgtttaca 1192
D 601 GGCATGCCAGCCTTGTGCCAGTCTCTTGTGCTGCAAAATATAGACCATACGGTGTTTACA 660
QY 1193 gagattccccgagacccttgttatacataataatttagcacatcaggggtgtgagacctcaa 1252
D 661 GAGATCCCCGACGACCCCTTGTATACATAATTTAGCACATCAGGGGTGTGGAGCCTGCAA 720
QY 1253 gtacatatcctgactgggattgctcctgaatggtatggagctttagaaatgggtatttc 1312

```

```

Db 721 GTACATATCCTGATCTGGGATTCCTCTGAATGGTATGAGAGCTTTAGAAATGGGTATTTC 780
QY 1313 cagaatgggcaagagagcatgcccttgacaaggtgagcagtgtaacttttgaaggag 1372
D 781 CAGAATGGGCAAGAGGAGCATGCCCTTGACAAGGTCAGGACAGTTAACTTTTGAAGGAG 840
QY 1373 cagttgtacagcagatcgagattgtgacgcagtcaggggtatttcatctgggaggtcacaa 1432
D 841 CAGTTGTGACAGCAGATCGAATTGTGACCGTCAGTCAGGGTTATTCTGGAGGTCACAA 900
QY 1433 ctctgaaggtgacagggccctcaatgagctcttaagctcccgaaaaagtgattgaatg 1492
D 901 CTGCTGAAGTGGACAGAGGCGCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGTATTGAATG 960
QY 1493 gaattgaaatggaattgacattaatgattggaaacccccaccacacagacagtgctccctc 1552
D 961 GAATTGTAATGGAATTGACATTAATGATTGGAACCCACACAGACAAAGTGTCTCCCTC 1020
QY 1553 atcattattctgctgacgacctctctggaaagccaaatgtaaaagctgaattgcagaagg 1612
D 1021 ATCATTATTCTGTCGATGACCTCTCTGGAAGGCCAAATGTAAAGCTGAATTGCAGAAG 1080
QY 1613 agttgggtttcaacctgtaaggagagattgtcctctctgattggctttattggaagactggatt 1672
D 1081 AGTTGGGTTTACCTGTAAAGGAGGATGTTCTCTGATTGGCTTTATTGGAAGACTGGATT 1140
QY 1673 accagaagagcattgatctcattaaatggccattccagagctcatgagggagagcgtgc 1732
D 1141 ACCAGAAGGCAATTGATCTCATTTAAATGGCCATTTCCAGAGCTCATGAGGAGGAGCTGC 1200
QY 1733 aatttgtcatgctggtgattggggatcccaatttttgaagctgagatgagatctaccaggt 1792
D 1201 AATTGTGTCATGCTGGATCTGGGATCCCAATTTTGAAGGCTGATGAGATCTACCGAGT 1260
QY 1793 csggtlacaaagataaattccgtggatgggttggtattgatttccagtttccacacagaa 1852
D 1261 CGAGTTACAAAGGATAAATTCCTGGATGGGTTGGATTTAGTTTCCAGTTTCCACAGAA 1320
QY 1853 taactgcaggttcgcgatatattgttaagccatcgagattgaaaccttgcggcttaactc 1912
D 1321 TAATCGAGGTTGCGATATATTGTTAATGCCATCGAATTTGAACCTTTCGGGCTTTAATC 1380
QY 1913 agctatgctatgcaataggtcacagttcctgtagttcatggaactgggggctccgag 1972
D 1381 AGCTATATGTTATGCAATATGTTACAGTTCTCTGTAGTTTCATGGAACCTGGGGCTCCGAG 1440
QY 1973 acagtcgagacacttcaaccttttgggtgcaaaagagaggggttacagggttgggcgt 2032
D 1441 ACAGTCGAGAGCCTTCAACCCCTTTTGGTGCAAAAGGAGAGAGGGTACAGGGTGGCGCT 1500
QY 2033 tctcacgcgttaacgctgacagaagtgttgggcatgtgcgaacgcgagatcgacattca 2092
D 1501 TCTCACGCTTAACCGTGGCAAGATGTTGTGGCATTTGCCAACCCTTCGTCGACATTCA 1560
QY 2093 ggagacacaagcgtcctggggagggctcatgaagcagagcagatgacgaagaccatcgt 2152
D 1561 GGGAGCACAGCCGCTCTCGGAGGGGCTCATGAAGCCAGGCATGACCAAGAACCATACGT 1620
QY 2153 gggaccatgccccgagcagtacgagcagatcttcgagtggggcttcgtggaacaccccta 2212
D 1621 GGGACCATGCCCGAGCAGTACGAGCAGATCTTCCAGTGGGGCTTTCGTGGACCAACCCCTA 1680
QY 2213 cgtcatgtagacggggagctggggaggtcccaagtcgaggtcctcctcagctcgaagacat 2272
D 1681 CGTCATGTAGACGGGGACTGGGAGGTCCCAAGTGCAGTCTCCTTCAGCTCTGAAGACAT 1740
QY 2273 cctcttcactctccgcgcgcggaaggaataccctctgacattgctgttctcctgctacag 2332
D 1741 CCTCTTCATCTTCCGCGCGCCCGGAAAGATACCCCTGTACATTCGCTGTCTCTGCTACAG 1800
QY 2333 tagagtcgcaatgcgcctgcttcttgggttcgccggttccgagaacata tgaaggctgtg 2392
D 1801 TAGAGTCGCAATGCGCCTGCTTGTCTTGGTTGCGCGCTTCGAGAACATATGACGCGCTGTG 1860

```



```

Db 810 AATGGGAGGATATATTATGGACAGAAATTCGATGTTTGTGTCAATGATGGCATCCCA 869
Qy 1142 gcttggcagagtccttctgctgcaaaatagaccatcggtgtttacagagattccc 1201
Db 870 GTCAGTGGCCAGTCTCTTCTGCTGCAAAATATAGACCATATGGTGTATTAAGACTCCC 929
Qy 1202 gacagacccttggatatacaataattagcatcatcagggtgtggagccctgcaagtacatc 1261
Db 930 GCACGATTCCTGTAATACATAAATTTAGCACATAGGGGTGAGAGCCCTGCAAGCACATATC 989
Qy 1262 ctgactcgggaatgcctcctcctgaaatgggtatggagcttttagaatgggtattccagaaatggg 1321
Db 990 CTGACCTTGGGTGGCCACCTGAAATGGTATGGAGCTCTGGAGTGGGTATTCCTCGAATGGG 1049
Qy 1322 caaaggagcatgcccctgacaaaggtcagcgagtttaacttttgaagagcaggttgtga 1381
Db 1050 CGAGAGGCATGCCCTCTGCAGAAAGGCTGAGGCAGTAAATTTTGAAGGTGCAGTTGTGA 1109
Qy 1382 cagcagatcggaattgtgacgctcagtcaggtgtattcatggaggtcacaaactgctgaag 1441
Db 1110 CAGCAGATCGAATCGTACTCTCAGTAAAGGTTTATTCATGGAGGTGACAACTGCTGAAG 1169
Qy 1442 gtgacaggggctcaatgagctcttaagctcccgaaagtgatgtattgaatggaattgtaa 1501
Db 1170 GTGCACAGGGGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATTAACGGAAATTTGTA 1229
Qy 1502 atgaattgacataatgattggaagccaccaccagacagagtgctcctcatcatatt 1561
Db 1230 ATGAATGACATTAATGATTGGAACCCCTGCCACAGACAAATGTATCCCTGCTCATATT 1289
Qy 1562 ctgctgagtcctctctggaagccaaatgaaagctgaaattcagaagaggttggtt 1621
Db 1290 CTGTTGATGACCTCTCTGGAAGGCCAAATGCTAAAGGTGCTATTCGAGAAGGAGCTGGTT 1349
Qy 1622 tacctgtaaggagagatgttctctgattggtcttattggaagactggattaccagaaag 1681
Db 1350 TACCTATAAGGCCCTGATGTTCTCTGATTTGCTTATTTGGAAGATTGGATTATCAGAAAG 1409
Qy 1682 gcattgactctataaaatggccattccagagctcatgagggagagcgtagctgttca 1741
Db 1410 GCATTGATCTCATCAACTTATACCATACCATGCTCATGCGGGAAGATTTCATTAATTGCTCA 1469
Qy 1742 tgcctggatctcgggagatcccaatttttgaaggtcgatgagatctacccagagtcgagttaca 1801
Db 1470 TGCTTGGATCTGGTGACCGAGCTTGAAGATTGGATGGATGATGATGATGATGATGATGATGAT 1529
Qy 1802 aggataaattcggtaggtgtgattttagtttagtttagtttagtttagtttagtttagtttag 1861
Db 1530 AGGATAAATTTGCGTGGATGGGTGGATTTAGTGTTCAGTTTCCAGTTTCCACCGAATAACTGCCG 1589
Qy 1862 gttgcgatatattgttaatgcatcgagatttgaaccttgcggttcttaatacagctatatg 1921
Db 1590 GCTGGCATATATTGTTAATGCCATCCAGATTGCAACCTTGTGGTCTCAATCAGCTATATG 1649
Qy 1922 ctatgcaataatggtacagttcctctgattttagtttagtttagtttagtttagtttagtttag 1981
Db 1650 CTATGCAATGAGGAGGCTTCTGTTGCTCATGCAACTGCGGGGCTTACAGATACCGGTG 1709
Qy 1982 agaccttcaaccttttggtaaaagagagagaggttacagggtggggtgttctcacccg 2041
Db 1710 AGAAGTTCAACCCCTTTCGTTGAGAAATGAGAGAGAGGTTACAGGGTGGGCAATTCGACCCC 1769
Qy 2042 taacctgtgacaagaatgtgtggcattgcgaaccgcgagtgctgacattcagggagcaca 2101
Db 1770 TAACACAGAAACATGTTGGTGGATTTGCGAATTCGCAACTGCAATATCATACAGGGGAACACA 1829
Qy 2102 agcgtcctcgggaggggtcctatgaagcagagcagatgacagaaagaccatactgtggaccatg 2161
Db 1830 AGTCTCTCTGGGAGGGGCTAATGAAGCGAGGCGATGTCAAAAAGACTTTCACGTGGGACCATG 1889
Qy 2162 cc-cagacagatagacagatcttcgagtggtggtccttcgtgagcaaacctacgtcatgt 2220

```

```

Db 1890 CCGCTGAACAATACGAACAATCTTCAGTGGGCTTCATCGATCGACCCCTATGTCTATGT 1949
Qy 2221 agcggggagctggggaggt 2239
Db 1950 AAAAAAGGACCAAAAGTGT 1968
RESULT 3
US-08-572-951-1
: Sequence 1, Application us/08572951
: Patent No. 5824790
: GENERAL INFORMATION:
: APPLICANT: KEELING, PETER L.
: APPLICANT: KNIGHT, MARY E.
: APPLICANT: GUAN, HANPING
: TITLE OF INVENTION: MODIFICATION OF STARCH
: TITLE OF INVENTION: SYNTHESIS IN PLANTS
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN DABY & CUSHMAN
: ADDRESSEE: Intellectual Property Group of
: ADDRESSEE: Pillsbury Madison & Sutro LLP
: STREET: 1100 New York Avenue, N.W.
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3918
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/572,951
: FILING DATE: 15-DEC-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/346,602
: FILING DATE: 29-NOV-1994
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/263,921
: FILING DATE: 21-JUN-1994
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Paul N. Kokulis
: REGISTRATION NUMBER: 16,773
: REFERENCE/DOCKET NUMBER: 222957/1.02.15C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 861-3000
: TELEFAX: (202) 822-0944
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2990 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: US-08-572-951-1

```

```

Query Match 48.1%; Score 1332; DB 1; Length 2990;
Best Local Similarity 81.7%; Pred. No. 1.8e-277;
Matches 1617; Conservative 0; Mismatches 340; Indels 21; Gaps 6;
Qy 237 cgaacctcctccgctcctccctgtcccgcgccacacccatggcgagcgggcggt 296
Db 537 GCACATGTTGCGGCGATCCACGGCCCCCCCCCGGCTCTGTCTCTCCCTCT 596
Qy 297 cggcgcggtgtcctcgcgccccagctccgcctcgcgcgatccgcgagcgcgggccg 356
Db 597 CCGCAATGGCAGCGCCCTCGGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 656

```


Qy	2088	attcaggggacacaagccgctcgtggaggggcctctgaagcgaagccatgacgaaagccca	2147
Db <td>1672</td> <td>ATACGGGGACACAAGTCCTCTGGGAAGGGCTAATGAAGCGAGGCATGTCAAAAGACTT</td> <td>1731</td>	1672	ATACGGGGACACAAGTCCTCTGGGAAGGGCTAATGAAGCGAGGCATGTCAAAAGACTT	1731
Qy <td>2148</td> <td>tacgtgggaccatgcccc</td> <td>2165</td>	2148	tacgtgggaccatgcccc	2165
Db <td>1732</td> <td>CACGTGGGACCATGGCCG</td> <td>1749</td>	1732	CACGTGGGACCATGGCCG	1749

```

RESULT      5
US-08-941-445A-20
: Sequence 20, Application US/08941445A
: Patent No. 6107060
: GENERAL INFORMATION:
: APPLICANT: Keeling, Peter
: APPLICANT: Guan, Haining
: TITLE OF INVENTION: Starch Encapsulation
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Circle
:

```

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
:

```

PRIOR APPLICATION DATA: US 60/026,855
 APPLICATION NUMBER: 30-SEP-1996
 FILING DATE: ATTORNEY/AGENT INFORMATION:
 NAME: Winner, Ellen P
 REGISTRATION NUMBER: 28,547
 REFERENCE/DOCKET NUMBER: 89-97
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1620 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: not relevant
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1620
 US-08-941-445A-20

Query Match	45.1%	Score 1248.4;	DB 3;	Length 1620;
Best Local Similarity	86.2%;	Pred. NO. 1.3e-259;		
Matches 1392;	Conservative	0;	Mismatches 221;	Indels 11;
Gaps 1;				

Qy	553	gacgcggcgctgggggaactcgcgcgcgcacctcctcctcgaaggagattcgtcgaagattcc	612
Db	4	gtccgcgagcttgacgcggagagacctcggcttcgaacctgaaggagattcctgaagtccc	63
Qy	613	atcgacgagcataattgtggctgcgaatgagcaggattctgagatcatcgatgcgaatgag	672
Db	64	atcgataaacacagtagttgtggcaagtagacacaagattcttgagattctgttgcgaaggag	123
Qy	673	caacctcaagctaagaattcacgtagcatcgtttgtgactggtgaaagctcctctat	732
Db	124	caagctcgagcttaagtaaacacaagcattgtcttggtaacgcgcgaagcttctcttat	183

Qy	733	gcaaa	gtcagg	ggggg	ttggg	agatg	tttgt	gtgtt	tcg	ttacc	aaattg	tctct	gtgtg	ctcgt	732																
Db	184	GCAA	AGTCTGGG	GGTCTAG	GAGATGTTT	TGGTGTTCAT	TGCCAGT	TGCTCTGT	GTGCTGT	CGT	TCGTG	CTGTG	CTGTG	CTGT	243																
Qy	793	ggtc	cacg	agtg	atg	ttg	taatg	cca	agata	ctt	aaatg	gg	tcc	ctg	ata	aaac	ctat	852													
Db	244	GGT	CACCGT	GTATGGT	TGTAT	ATGCC	AGATATTT	TAAT	ATG	TACT	CGT	ACT	CCG	ATA	TAGA	ATTAT	303														
Qy	853	gcaaa	gcat	tata	cactg	cgaa	gcac	at	taaga	t	ccatt	gctt	gtt	ggg	gat	ccac	at	912													
Db	304	GCAA	ATG	CAATTTT	TAC	ACAGAAA	ACATTC	CGAT	TCCAT	GCTTT	TGG	CGGT	TGA	CAAT	GA	ATGA	363														
Qy	913	gtg	ac	ctttt	t	cat	ag	tat	aga	ca	aac	gt	cat	gg	tgt	ttg	cg	at	cc	gt	ca	972									
Db	364	GTT	ACCT	TTCTT	CCAT	GAG	TAT	AG	ATTC	CAG	TGT	ACT	GGG	TGT	TTG	TAT	CAT	CC	CT	CA	423										
Qy	973	tat	ca	gac	ac	gag	ga	gtt	tata	t	gag	ata	aa	ttt	gt	g	ct	ttt	gt	g	t	at	ca	1032							
Db	424	TAT	CAC	AG	CTCG	AAATTTAT	TAT	GAG	ATA	AGTTT	TG	CTTTT	TG	TG	TG	TG	TG	TG	TG	TG	TG	TG	TG	483							
Qy	1033	agata	ca	ac	ac	ct	ctt	gt	ctat	g	ctg	cat	g	cg	ag	cc	ca	ct	aa	ct	ctt	g	at	tg	ga	ga	1092				
Db	484	AGAT	PAC	ACAT	CTCT	TTT	GCTAT	G	CTAT	G	CTAT	G	AGG	CTCT	T	GTAT	T	CC	TT	T	G	AA	T	T	G	G	G	543			
Qy	1093	tata	tt	at	tg	gac	aga	at	tg	cat	gtt	gtt	gt	g	a	ac	ga	t	tg	g	cat	g	cc	ag	ct	gt	g	cc	1152		
Db	544	TAT	ATTTT	TAT	GC	ACAG	AA	TTCAT	T	GT	T	GT	T	GT	T	GT	T	GT	T	GT	T	GT	T	GT	T	GT	T	GT	603		
Qy	1153	gtc	ct	ct	gt	ct	g	ca	aa	tata	g	ac	cat	g	ct	g	tt	at	c	a	g	at	t	c	c	g	a	c	ct	1212	
Db	604	GT	CTCT	CTCT	GCT	GCA	AAATAT	AG	ACATAT	G	TG	TG	TTT	TAT	AA	AG	AC	T	C	C	G	A	G	A	T	T	C	T	663		
Qy	1213	gt	t	a	c	a	t	a	a	tt	t	a	c	a	t	g	g	a	c	ct	g	g	a	c	ct	g	g	a	c	1272	
Db	664	GTA	TAT	AT	AT	TTT	TAG	CAC	AT	C	AGG	GT	T	AG	A	G	CT	T	G	C	A	C	A	T	A	T	C	T	723		
Qy	1273	ttg	ct	ct	ct	g	a	at	g	g	at	g	g	g	t	tt	c	c	a	g	a	at	g	g	g	g	a	g	g	1332	
Db	724	TTG	CAC	CT	CAAT	TGGTAT	T	GAG	CT	C	T	G	A	G	T	TT	T	C	C	T	G	A	T	T	C	C	T	G	A	783	
Qy	1333	gcc	ct	t	a	c	a	ag	g	g	t	g	a	g	a	c	a	g	a	g	a	g	a	g	a	g	a	g	a	1392	
Db	784	GCC	CTT	G	CA	AGG	T	GAG	G	CAG	T	TAATTTTT	T	G	A	A	AG	T	G	C	A	G	T	T	G	T	G	C	A	843	
Qy	1393	at	t	g	a	c	g	t	c	a	g	g	t	a	t	t	c	a	a	c	t	g	t	a	a	g	g	t	g	a	1452
Db	844	AT	C	G	T	G	A	C	T	G	T	A	A	G	G	T	T	A	T	T	T	T	T	T	T	T	T	T	T	903	
Qy	1453	ctc	a	a	t	g	a	g	c	t	t	a	a	g	t	g	a	a	a	a	a										

```
Qy 1813 cgtgatgggttgatttagtgcagttccagttcccccacagaataactgcaggttgcatata 1872
Db 1264 CCGTGGTGGTGGGATTTAGTGTTCAGTTTCCACCGGAATAACTGCCGGCTGGCATATA 1323
Qy 1873 ttgttaatgccatgcagattgaacctgcggtcttaatacagctatgctatcaaat 1932
Db 1324 TTGTTAATGCCATGCAGATTGCAACCTTGCTGCTCAATCAGCTATATGCTATCAGTAT 1383
Qy 1933 ggtacagttcctgtagttcatggaactggggggtcccgagacacagtcagagaccttcaac 1992
Db 1384 GGCACAGTTCCTGTGTTCATGCAACTGGGGCCCTTAGAGATACCGTGGAGAACTTCAAC 1443
Qy 1993 ccttttggtgcaaaaggagagaggggtacaggggtgaggggttctcacgcgtacacgtggac 2052
Db 1444 CTTTCGGTGAAGATGGAGACAGGGTACAGGGTGGCGATTCGCACCCCTTAACACAGAA 1503
Qy 2053 aagatg-ttgtgggctatgcaacgcgagatgcacattcagggagcacagccgtcctg 2111
Db 1504 AACATGTTGTGGACATTGCCAACTGCATATCTACATACAGGGAACACAAAGTCTCCTG 1563
Qy 2112 ggaagggtcctatgaagcgaggtgacgaaagacacatacgtgggacctgcccc 2165
Db 1564 GCAAGGGCTAATGAAGGAGGAGCATGTCAAAAGACTTCACGTGGGACCATGCGCG 1617

RESULT 6
US-08-836-567-9
; Sequence 9, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Ag-revo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
```

```
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Desire
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in Lambda ZAPII
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 68..1990
; US-08-836-567-9
```

```
Query Match 27.2%; Score 752.8; DB 3; Length 2360;
Best Local Similarity 68.3%; Pred. No. 7e-153;
Matches 1088; Conservative 0; Mismatches 497; Indels 7; Gaps 3;
```

```
Qy 633 tgaagtgagcaggattcctgagatcattggtgcaatgagcaacctcaagctaaagtac 692
Db 397 TGATAGAGATGACTTGAGTGAGGATAGTGGAGAGATGGAGAAACCCCAATCAAAATTAAC 456
Qy 693 acgtagcatcgtgtttgtgactgtgaagctgctcccttgcgaagtcagaggggggttggg 752
Db 457 TTTCATATATCATTTTGTACTGCTGAAGCAGCTCCATATTTCTAAGACTGGTGGATTAGG 516
Qy 753 agatgtttgtggttcgtaccgaattgctctgtgctgctggtgcacgcagtgatggttgt 812
Db 517 AGATGTTGTGGTCTTTTGCAATGGCACATAGCTGCTCGGGGTCACTGCTAATGTCGT 576
Qy 813 aatgcaagatacttaaat---gggtcctctgataaaaaactatgcaaaagcatatcac 869
Db 577 TTCACCTAGGTATTTGAATGGAGGTCTTCAGATGAAAGTAGCGCAATGCTTTGACCT 636
Qy 870 tgggaagcacattaaagattccatgctttgggggatcacatgaaagtgcacctttttcatga 929
Db 637 TGATGTCGGGGCCACATGTCATTGCTTGGTGTGATGCACAGAGAGTACCCCTTCTACCATGA 696
Qy 930 gtatagagacaacgtcgattgggtgtttgtcgatcatccgtcatatcacagaccaggaag 989
Db 697 ATACAGGGCAGGTGTTGATGGGTATTTGTGGACCACTCTTCTACTGCAGACCTGGAAC 756
Qy 990 ttatatggagataatttgggtgcttttggatgaatacagttcagatcacacactccttg 1049
Db 757 GCCATATGGTGATATTTATGTCGATTTGGTGATATCAGTTTCGCTTTCATCTTGTCTTC 816
Qy 1050 ctatgctgcacgagggccccactaaaccttgaattggaggagatatattatggacagaa 1109
Db 817 TCACGCGAGCATGTGAAGCGCATTTGGTTCTTCCACTGGGAGGGTTCACTTATGAGAGAA 876
Qy 1110 ttgcatgtttgtgaaacgattggcatgccagccttgtgccagtcctcttctgtgcaaa 1169
Db 877 GTGCTTGTTCGCTAAATGATTTGGCATGCTGCCCTGGTTCCCTTTTAGCGGCCAA 936
Qy 1170 atatagaccatacgggtttacagagattccgcgcagcacaccttggtatacaataattagc 1229
Db 937 GTATCGTCTTATGGTGTTTACAAAGGATGCTCGTAGATTTGCGCAATACACACATTCG 996
Qy 1230 acatcaggggtgagacgtgcgaagtacatctctgattggtggtcctcctgaaatggtca 1289
Db 997 ACATCAGGGAGTGGAGCCTGCAGTAACCTACAATAATTTGGGTTTGGCTCCCAATGTGA 1056
Qy 1290 tggagctttagaatgggtatttccagaatgggcaagagagcatgcccttgacagaaggtga 1349
Db 1057 TGGAGCAGTTGAATGATATTTCCCAATGGGCAAGGGCGCATGGCTTGTGACACTGGTGA 1116
Qy 1350 ggcagttaacttttgaaggagcagttgtgacagcagatcggaattgtgacgcgtcagta 1409
Db 1117 AACAGTGAACGTTTTCGAAAGGGCAATAGCAGTTGCTGATCGATACGATACAGATTAGCCA 1176
Qy 1410 ggggtattcatgggaggttcacacactgctgaaggtgagacagggccctcaatgagctcttaag 1469
Db 1177 GGGATACTCATGGGAATAAACAACTCCTGAAGGGGATATGGGCTACATGAGCTGTTGAG 1236
Qy 1470 cttccgaaaaaagtgtattgaaatgtgaaatggaattgacattaatgattggaaccc 1529
```

```
Db 1237 CAGTACACAGTCTGTCTTCTTAATGGAATTAATAATGGAATAGATGTTAATGATTTGGAACCC 1296
Qy 1530 caccacagagaagtgtctccctcaccattctctgtcgtacacctctctggaaagccaa 1589
Db 1297 GTGCGACAGATGACATATTTGCTTGGCATTTACCTCAATGACCTCTCCGAAAGGTGCA 1356
Qy 1590 atgtaagcgtgaattgcagaagagttgggtttacctgtlaagggagagtgcttccctgat 1649
Db 1357 GTGCAAGACTGATCTGCAAAAGGAAGTGGGCTTCCAATTCGACCTGATTTGCTCTGAT 1416
Qy 1650 tggctttattgaaagactgattaccagaagagcattgtatctcattaaatggccattcc 1709
Db 1417 TGATTATTGGAAGGCTGGACTTACCAGAAAGGTGTTGACATAATCTCTGCAGCAATTCC 1476
Qy 1710 agagctcatgaggaggagtgatctgtcgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1769
Db 1477 AGAAGTATGACAGATGATGCTCAAGTGTATGCTTGGATCTGGTGAGAAACAATATGA 1536
Qy 1770 aggctggatgagatctaccgagtcgagttacaagataaattccgtggatgggttgatt 1829
Db 1537 AGACTGGATGACATACAGAAATCTTTTAAAGACAAATTTCTGCTTGGTTGGATT 1596
Qy 1830 tagttccagtttcccaagataactgcaggttgcgtatattgttaattgccatcgag 1889
Db 1597 TAATGTTCCAGTTTCTCATAGGATAACAGCAGGATGCGACATACACTATTGATGCCCTCAAG 1656
Qy 1890 attgaaccttcggtcttaacatcagctatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1949
Db 1657 ATTCGACCGTGTGGCTTAACCAATTTGATGCAATGAGATATGACCATACCTATTGTT 1716
Qy 1950 tcatggaactggggcctccgagacacagtcgagaccttcaaccttttgggtgcgaagg 2009
Db 1717 TCATAGCAGGGGGGCTAAGAGACACAGTGAAGATTTTAATCCATATGCTCAAGAGG 1776
Qy 2010 ---agaggaggtacaggggtggggtctcaccgttaaccgttggacagagatgttggtgc 2066
Db 1777 ATAGGTGAAGTACGGGTGACATTTTCTCTCTAAGCAGTGAAGAAAGTTGCTTGATAC 1836
Qy 2067 attgcgaaccgcatgtcagattcagggagcacagcgctcctcgtggaggggctcatgaa 2126
Db 1837 ACTGAAGTGGCAATCGGACTTATACAGACATAGTCATCTTGGAGGGGATTGATGAG 1896
Qy 2127 gcgagcagacgaagaccatcacgtgggaccatgc-cccagcagcagcagcagcagcattt 2185
Db 1897 CAGAGGTATGGGAAGGACTATTCTGGGAAATGCAGCATTCAATATGAACAAGTTT 1956
Qy 2186 cgaagtgcccttcgtggaccacacctacgtca 2217
Db 1957 CACCTGGGCTTTATAGATCCTCCATATGTCA 1988
```

```
RESULT 7
US-08-836-567-3
; Sequence 3, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1758 base pairs
; TYPE: nucleotide
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Berolina
; TISSUE TYPE: tuber tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in pBluescriptSKII+
; FEATURE:
; NAME/KEY: CDS /function= "Polymerization of
; LOCATION: 1..1377 starch"
; OTHER INFORMATION:
; OTHER INFORMATION: /product= "Starch synthase"
; OTHER INFORMATION:
US-08-836-567-3
```

```
Query Match 23.4%; Score 648.4; DB 3; Length 1758;
Best Local Similarity 68.6%; Pred. No. 1.8e-130;
Matches 941; Conservative 0; Mismatches 421; Indels 10; Gaps 3;

Qy 856 aaggcattatcacactgcgaagcacatttaagattcccatgctttgggggacacacatgaagtg 915
Db 10 AATGCTGTGACCTTGATGCGGGCCACTGTCCATTGCTTTGGTGATGCACAGGAAGTA 69
Qy 916 acccttttcatgagtagagacaacacgtcgtgattgggtgtttgtcgcatacctcgtcat 975
Db 70 GCCTTCTACCATGAATACAGGCGCAGGTGTGATTGGGTATTGTGGACACACTCTCTTAC 129
Qy 976 cacagaccagaagtttatatggagaataatttgggtgcttttgggtgataaatacagtcaga 1035
Db 130 CCGAGCCTGGAAACCCCATATGATATTTATGCTGATGCTTTGGTGAATAACAGTTTCG 189
Qy 1036 tacacactccttggctatgctgcagcgagcccccactaaaccttgaattggaggagat 1095
Db 190 TTCACCTTTGCTTCTCACGCGACGATGTAAGCGCCATTGGTCTTCCACTGGAGGGTTC 249
Qy 1096 atttatgacagaattgcatgttttgggtgaaacgattggcagcgttggccagtc 1155
Db 250 ACTTATGGAGAGAGTGTCTTCTCGCTAATGATTGCAACGCTGCCTTGGTCTCTTTA 309
Qy 1156 cttcttgcgcataataatagaccatacagtggttttacagagattcccgagcagaccctgtt 1215
Db 310 CTTTATGGCGCCAGTAGTATGCTCTTATGTTTAAAGGATGCTCGTAGTATTGTCGA 369
Qy 1216 atacataatttagcacatcagggtgtggagcgtgcaagtacatatacctgatctgggattg 1275
```

```

Db 370 ATACACAACATTGCACATCAGGAGTGGAGCCTGCGAGTAACTACATATATTTGGGTTG 429
QY 1276 cctccgaatggtatgagctttagaattggtatttccagaatggcgaagggagcagtcgc 1335
Db 430 CTCCCAATGTTGATGGAGCTGTAATGATATTTCCCATGTCGCGCAGGGCGCATGG 489
QY 1336 cttgacaagggtgagcagcttaacttttgaaggagcagctgtgtgacagcagatcgatt 1395
Db 490 CTTGACACTGGTGAACACAGTGAACGTTTGAAGGGGCAATAGCAGTTGCTGATCGGATA 549
QY 1396 gtgaacgctcagtcaggggttatcattgggaggttcacaacgctgaaaggttggaagggcctc 1455
Db 550 CTGACAGTTAGCCAGGATGACTCATGGGAATAAACAACTCTCGAAGGGGATATGGGCTA 609
QY 1456 aatgagctcttaagctcccgaaaggtatgaatggaattggaattgaattgaattgacatt 1515
Db 610 CATGACCTGTTGACGAGTAGACAGCTGTTCTTAATGGAATTAATGAATGATGATGTT 669
QY 1516 aatgattggaaccccccacagacaagtgctccctcattcatttctgtcgtatgacctc 1575
Db 670 AATGATTGGAACCGTCGACAGATGAGCATATCGCTTCGCAATTACTTCCATCAATGACCTC 729
QY 1576 t-----ctggaagggcccaaatgtaagctgaattcagaagaggttggtttacctgta 1629
Db 730 TCCCCCCTGGAAAGGTTGAGTCAAGACTGATCTGCAAAAGGAAGTGGGCGCTTCCCAAT 789
QY 1630 agggagagtgcttctctgattggtcttattggaagctggtattaccagaagsgcattgat 1689
Db 790 CGACCCGATGTCCTGATGATTGATTATGGAAGCTGGACTACCAAGAGGTGTGAC 849
QY 1690 cteattaaatggccattccagagctcatgagggaggaagctgcaatttgcgtgctgga 1749
Db 850 ATATCTGTCGACAAATTCAGAACTATGCAAGATGATGTCGAAGTTGTAATGCTTGA 909
QY 1750 tctgggagatcccaatttttaagctgagatgagatctaccagagtcgagtcgatttacaaggataa 1809
Db 910 TCTGGTGAGNAACAATGATGAAGACTGGATGAGACATACAGAAATCTTTTAAAGACAAA 969
QY 1810 ttcctggatgggttgatttagtttccagtttccacagacaagaactcaggttgcgat 1869
Db 970 TTTCTGCTGGGTTGGATTTAATGTTCCAGTTTCTCATAGGATAACAGCAGGATCGAC 1029
QY 1870 atattgttaagtcagcagagattgaaccttgcggtctttaaatacagctatatgctatgcaa 1929
Db 1030 ATACTATTGATGCCCTCAAGATTTCGAACCGTGTGGCTTAAACCAATTTGATGCAATGAGA 1089
QY 1930 tatggtacagttctctgattgctgaactgggggctccgagacacagtcgagaccttc 1989
Db 1090 TATGGCACCATACCTATTGTTTCATAGCACGGGGGGCTTANGACACACAGTGAAGGATTTT 1149
QY 1990 aacccttttggtgcaaaagg---agaggaggttacagggtggcggttctcaccgcctaacc 2046
Db 1150 AATCCATATGCTCAAGAAAGGAAAGGTGAAGGTACCGGTGGACATTTTCTCCTCAAGC 1209
QY 2047 gtgacaagatgtgtggtgcatgcaacgcgcatgtcagattcaggaggaagcaagccg 2106
Db 1210 AGTGAAGAGTTGTTGATACACTGAAAGTGGCGATCAGGACTTATACAGAACATAAGTCA 1269
QY 2107 tcttgaggaggggctcatgaagcagagcagatgacgaagaccatacgtgggaccatgc-ccc 2165
Db 1270 TCTTGGGAGGATTTGATGAAGAGAGGTATGGAAGGGGACTATTTCTTGGGNAATGACGC 1329
QY 2166 gagcagtaagcagatctctcagagtgggccttcgttggaaccaaccctacgtca 2217
Db 1330 ATTCAATATGACGAAGTTTTCACCTGGCGCTTTATAGATCCTCCATATGTCA 1381

```

RESULT 8

US-08-572-951-3
; Sequence 3, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:

```

; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: MODIFICATION OF STARCH
; TITLE OF INVENTION: SYNTHESIS IN PLANTS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: Pillsbury Madison & Sutro LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,951
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,602
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/263,921
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul N. Kokulis
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222957/1.02.15C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2380 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-572-951-3

```

```

Query Match 12.6%; Score 349; DB 1; Length 2380;
Best Local Similarity 39.2%; Pred. No. 5.6e-66;
Matches 583; Conservative 220; Mismatches 609; Indels 75; Gaps 7;

QY 697 agcatcggtgttgactggtgaagctgctcccttatgcaaaagtcaggggggttggagat 756
Db 679 AAYGTGNTGNTGTCNGCNWSNGARTGYCCNCTTTGYAARACNGNGNGNYNGNGAY 738
QY 757 gtttgggttcgttaccattgctcttgcgtcgtggtcaccgagtgatggttgtaaty 816
Db 739 GTNGTNGGNCNVTNCCNAARGCNYTNGCNMNGNGNCAYMNGNTATGTTGTTNGT 798
QY 817 ccaagataacttaattgggtgctcctctgataaaaaactatgcaaaagcattatcacactgcgaag 876
Db 799 CCNMNGNTAYGNGNGAR-----TAYGCNGARGCNMNGNGAYTNGNGNTN 840
QY 877 cacattaaagattccatgcttctgggggagtcacatgaagtgacctttttcatgagtaga 936
Db 841 MGNMNGNTAYAARGTNGCNGCNGCARGAWSNGARGTACNTATTTTCAYWNTAYATH 900
QY 937 gacaacgtcgatgggtgtttgttcgcatccctgcattatcacagacaggaagtatat 996
Db 901 GAYGGNCTNGAYTGTGTTNTTGTNGARGCNCNCCNTTY-----MGNCAYMGNCAAYAY 954

```


[illegible]

```

Db 2002 ATHGAYCNYTNWNCAYTGYYTNACNACNTAYMGNAAYTAYAAARGAWSNNTGGMNGCN 2061
Oy 2119 ctcatgaagcagcagcatgcaagaagaccatagctgggaccatgcgcc 2165
Db 2062 TGYMGCNCGNMGNGTGGCGARGAYYTNWNTGGAYCAYGCGNC 2108

RESULT 9
US-08-836-567-7
; Sequence 7, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRA
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2793 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv Desiree
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cdna-library in Lambda ZAPII
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 242..2542
US-08-836-567-7

```

Query Match 10.4%; Score 287; DB 3; Length 2793;
Best Local Similarity 53.9%; Pred. No. 1.3e-52;
Matches 801: Conservative 0; Mismatches 605; Indels 8

Qy 697 agcatcgtgttttgtgactggtgaagctgctccttatgcaaaagtcaggggggttgaggagat 756

Db	1070	ACACATTAATTTGGTGGCTTCAGATCCGCTCCATGGTCTAAACACAGGTGGCGCTTGGAGAT	1129
Qy	757	gtttggttctgtaaccaattctctctgctctgctggtccaccagtgatggttgtaagt	816
Db	1130	GTTGCTGGAGCAATTACCCAAAGCTTTGGCTGCAGCTGGCCACAGATTATGTTGTTGGCA	1189
Qy	817	ccaagatacttaaatgggtccctctgataaaaaactatgcaaaaggcaatatacacctgcgaag	876
Db	1190	CTCTGGTT-----ATGACAACATATCTGAACTCAAGATTCTTGGTGTA	1231
Qy	877	cacattaagattccatgctttgtggggatcacatgaatgaatgacctttttcatgagtataga	936
Db	1232	AGAAAAATTTATAAGTTGATGGTCAGGATGTGGAAGTGACTTACTTCCAAGCTTTTATT	1291
Qy	937	gacaaagtcgattgggtgtttgtcga---tcactcgtcatatcacacgaacggaagtta	993
Db	1292	GATGGGTGGGATTTGCTTTTTCATTGACAGTCATATGTTTATAGACACATTGGGAACAACATT	1351
Qy	994	tatggagataaatttgggtccttttgggtataatcaatcagttcagatcacacactccttgcctat	1053
Db	1352	TACGGAGGGAACCGTCTGGATATTATAACGCATGGTTTAT-----TTTGGCAA	1402
Qy	1054	gctgcatacgagggccccaataatccttgaattgaggagatatattatggaca---gaat	1110
Db	1403	GCAGCCATTGAGGTTCTCTTGGCATGTTCCATGTGTGGGGTCTGCTATGGAGATGGAAAT	1462
Qy	1111	tgcatgtttgtgtgaacgattggcatccagccttgtgccagtccttcttgcctgcgaata	1170
Db	1463	TTAGTGTTCATTGCTAATGATTGGCAATACCTGCTTTATTGCCAGATATATCTGAAAGCTTAT	1522
Qy	1171	tatagaccatacgggtgtttcacagagattccgcagcacaccttggttatacataatttagca	1230
Db	1523	TATCGTGACAAATGGAATTAATGAACATACAAGATCTGCTGGTGATTATACATACATCGCT	1582
Qy	1231	catcaaggggtggagcctcgaagtacatatcctgacttgggattgcctcctgaaatgdat	1290
Db	1583	CATCAGGGTCGTGGTCCCTTTGGAGGATTTTCATATGTAGATCTTCCACACACTATPATG	1642
Qy	1291	ggagccttagaattgggtatttccagaatgggcaagaggcagtcaccttgacaagggctgag	1350
Db	1643	GACCCTTTCAAGTTGTATGACCCAGTAGGA-----GGTCAG	1678
Qy	1351	gcagttaaacttttgaagaggcagtttgacagcagatcggaattgtgacctcagtcag	1410
Db	1679	CATTTTCAACATATTTTGGCGCTGGCTTAAGACACAGACATCGTGTAGTTACAGTTAGTCAT	1738
Qy	1411	ggttattcctggaggtcacaaactcgaagtggagaggccctcaatgagctcttaagc	1470
Db	1739	GGATATTATGGGAATAAGACTTCCCAAGGTGGTGGGATTGCAATCAGATAAATTAAT	1798
Qy	1471	tcccgaagaagtgtattgaaatggaattgaaatggaattgacattaatgattggaacccc	1530
Db	1799	GAGAACGATTGGAAATTTACAGGGTATTGTGAATGGGATTGATACAAAGAGTGGAAACCT	1858
Qy	1531	accacagacaagtgtctccctcaatattctgtcgatgacctt-----	1576
Db	1859	GAGTTGGACGTTCACTTACAGTCACATGGTTACATGAACACTACTCCTTGGACACGCTACAG	1918
Qy	1577	-ctggaaagggccaaatgtaaagtgaattgacagaaggattgggtttacctgtaaagggag	1635
Db	1919	ACTGGCAAGCCCTCAATGTAAAGCTCAATTGGCAGAAGGAACCTTGCTTTTCCAGGTTCTGTAT	1978
Qy	1636	gatgttctctgatttggctttatttgggaagcgtgattaccagaaggcatatgatctcaatt	1695
Db	1979	GATGTCCCACTACTCGGTTTCATTGGGAGGCTTGACCCACAAAAAGGCGTTGATCTGATT	2038
Qy	1696	aaaaatggccattccagagctcatagggagagcagctgcaatttgcactgcttgatctggg	1755
Db	2039	GCTGAGGCCAGTGCTTTGGATGATGGGTACAGATGTACAACGTGTCATGTTGGGGACGGGG	2098
Qy	1756	gatccaatttttgaaggctgtgtagatgagatctaccagctcgagttcacaggagataatccctg	1815

Db 2099 AGGCGTGACCTTGACAGAGATGCTAAGGCAATTTGAGTGTCAACAATGATAAAATTAGA 2158

Qy 1816 ggaatgggttgattagttgtccagtttccacagaataaactgcaggttgcgatatattg 1875

Db 2159 GGATGGGTTGGTTCTCTGTGAAGACATCTCATCGTATAACTGCTGCTGCAGACATTTCTG 2218

Qy 1876 ttaatgcatacgagatttgaaccttgcgggtcttaatacagctatataatgcataataggt 1935

Db 2219 CTATGCGCTTCTAGATTGAGCCTTGGGGACTGAACAGACGCTTTATGCAATGAATATCGG 2278

Qy 1936 acagtctcgttagttcatggaactggggcctccgagacacacagtcgagaccttcaaccct 1995

Db 2279 ACTATTCTGTTGTCATGCAGTAGGAGGACTCAGACATACTGTCGACAGCCCTTTGATCCT 2338

Qy 1996 ttggtgcaaaagagagagggttacagggtggggcgttctcacccgttaaccgtggacaag 2055

Db 2339 TTTAATGAGTCAGGACTG-----GGGTGGACCTTCAGTAGGGCTGAAGCTAGCCAG 2389

Qy 2056 atgttggggcatgcgaaccgcgatgcacattcaggaagacacaaagccgtcctggag 2115

Db 2390 CTGATCCAGCGATTAGAAATTTGCTTACTGACTTATCTGAGTGATCAAAAAGAGTTGGGAG 2449

Qy 2116 ggggtcatgaagcaggcagtcacgaagacacatacgtgggaccatgc 2162

Db 2450 GGGATTACAGACACGTTGTATGACACAAGACTTAAGTTGGGATAATGC 2496

RESULT 10

US-08-836-567-5

Sequence 5, Application us/08836567

Patent No. 6130367

GENERAL INFORMATION:

APPLICANT: Kossmann, Jens

APPLICANT: Springer, Franziska

APPLICANT: Abel, Gernot

TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES

TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC

TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/836,567

FILING DATE: 24-JUL-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP95/04415

FILING DATE: 09-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 44 41 408.0

FILING DATE: 10-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: Agrido-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1926 base pairs

TYPE: nucleotide

STRANDEDNESS: unknown

:	TOPOLOGY:	linear
:	MOLECULE TYPE:	cDNA to mRNA
:	HYPOTHETICAL:	NO
:	ANTI-SENSE:	NO
:	ORIGINAL SOURCE:	
:	ORGANISM:	Solanum tuberosum
:	STRAIN:	cv. Berolina
:	TISSUE TYPE:	tuber tissue
:	IMMEDIATE SOURCE:	
:	LIBRARY:	cDNA-library in pBluescriptSK+
:	FEATURE:	
:	NAME/KEY:	CDS
:	LOCATION:	2..1675
:	OTHER INFORMATION:	/function= "Polymerization of
:	OTHER INFORMATION:	starch"
:	OTHER INFORMATION:	/product= "Starch synthase"
:	US-08-836-567-5	
	Query Match	10.1%; Score 279; DB 3; Length 1926;
	Best Local Similarity	53.5%; Pred. No. 5.8e-51;
	Matches	796; Conservative 0; Mismatches 610; Indels 81; Gaps
Qy	697	agcactcgctgtttgtgactggtagcgtgcctcttatgcaaaagtccaggggggttggagat 756
Db	203	AACATTATTTTGTGGCTTCAGAATCGCGTCCATGGTCTAAACAGGTTGGCCTTGGAAT 262
Qy	757	gtttgsgtttcgttaccaattcctcttgctcgtcgtggtccaccagtgatggttgtaatg 816
Db	263	GTTGCTGGAGCATTTACCCCAAGCTTTGGCTGCAGCTGCCACAGAGTTATGCTTGTGGCA 322
Qy	817	ccaagatacttaaatgggttcctctgatataaaaactatgcaaaaggcatatacacactgcgaag 876
Db	323	CCCTCGTT-----ATGACAACTATCTGAACTCAAGATTCCTGGTGTA 364
Qy	877	cacattaagattccatgctcttgggggataacatgaagtgcaccttttttcagtatataga 936
Db	365	AGAAAAATTTATAAAGTTGATGGCTCAGGATGTGGAAAGTACTTACTTCCAAGCTTTTATT 424
Qy	937	gacaaocgtcgattgggtgtttgtcga----tcatacgtcatatcacagacaggaaagtta 993
Db	425	GATGGTGTGGATTTTGTTCATTGCAGTCAATGTTTAGACACATYTGGAACAACATT 484
Qy	994	tatggagataattttggtgcttttggtgataatcagtttcagatacacactcctctgctat 1053
Db	485	TACGGAGGAACCGTGTGGATATTTTAAACGCATGTTTTAT-----TTTGCAA 535
Qy	1054	gctgcacgcgagccccactaactcttgaattggagagatatatttatggaca---gaat 1110
Db	536	GCAGCGATTGAGGTCCTTGGCATGTTCCATGTGGTGGGGTCTGCTATGGAGATGGAAAT 595
Qy	1111	tgcattgttctgtgaacgattggcatgccagccttctgcccagtcctcttctgctgcaaaa 1170
Db	596	TTAGTCTTCATTGCTAATGATTGGCATACTGTTTATTGGCCAGTATATCTGAAAAGCTTAT 655
Qy	1171	tatagaccatacggfitttacagagattcccgagcagacccttgttatacataatttagca 1230
Db	656	TATCTGCAANTGGAAATTATGAACATATAAAGATCTGCTCGGTGATTCTAACACATCGCT 715
Qy	1231	catcaggggtgtggagccctgcagtgatatactcctgatctgggattgctcctcctgaatggtat 1290
Db	716	CATCAGGGTCGTGGCTCTTTGGAGGATTTTTCATATGTAGATCTTCCACCACACTATATG 775
Qy	1291	ggagcttttagaatgggttattccagaatagggcaaggagccatgcccccttgacaagggtag 1350
Db	776	GACCCCTTCAAGTTGTATGACCCAGTAGGA-----GGTGAG 811
Qy	1351	gcagtttaactttttgaaaggagcagttgtgacagcagatcgattgtgacgctcagtcag 1410
Db	812	CATTTCAACATTTTTCGGCTGGCTTAAGACACAGCATCTGTAGTTACAGTTAGTCAT 871
Qy	1411	ggttattcatgggaggttcacaactgctgaaaggttgagagggccctcaatgagctcttaagc 1470

APPLICATION NUMBER: US/08/941.445A
 FILING DATE: 30-SEP-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,855
 FILING DATE: 30-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Wanner, Ellen P
 REGISTRATION NUMBER: 28,547
 REFERENCE/DOCKET NUMBER: 89-97
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2097 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: not relevant
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2097
 US-08-941-445A-10

Query Match	8.4%	Score 233	DB 3	Length 2097
Best Local Similarity	52.5%	Pred. No. 4.8e-41		
Matches 796	Conservative 0	Mismatches 645	Indels 76	Gaps 10
Qy	697	agcatcgctgttttgactggtgaagctgctccttatcgaagtcaggggggtttggagat	756	
Db	622	AACGTCGTCGGTGGCTTCGTAATGTGCTCCTTCTGCAAGACAGGTGGCTTGGAGAT	681	
Qy	757	gtttgtggttcgttaccaattgctctgctgctgctggttcacccagtgatggttaag	816	
Db	682	GTCTGGGTGTTTGGCTTAAGGCTCTGGCGAGGAGACACCGTGTATTGTCGTGATA	741	
Qy	817	ccaagatacttaaatgggtctctgataaaaaactatcgaagggcattatacactgcgaag	876	
Db	742	CCAAGATATGGA-----GAGTATGCCGAAGCCGGGATTTAGTGTGAAGGAGA	789	
Qy	877	cacattaaagattccatgcttttgggggacacatgaagtgacctttttcattgagtata	936	
Db	790	CGTTACAAAGGTAGCTGGACAGGATTCA-----GAAGTACTTATTTTCACTTC	843	
Qy	937	gacaacgtcgagattgggtgtttgtcgatcatccgttcataatcacagaccagggaagtttat	996	
Db	844	GATGGAGTTGATTTGTATTCGTAGAAGCCCTCCCT-TCGGGACCGGCAACAATAAT	902	
Qy	997	ggagataatttggcttttgggataaatcagttcagatatacacactcctttgctatgct	1056	
Db	903	-----TTATGGGGGAGAAGATTGGATATTTGAAGCCGATGATTTTGTCTGCAAGGCC	957	
Qy	1057	gcatacagggtcccccaataactccttgaaattggggagatataattatggaca---	1113	
Db	958	GCTGTTGAGGTTCCATGGTATGCTCCATGTGGCGGTACTGCTATATGGTATGGCAACTTA	1017	
Qy	1114	atgtttgttggaaacgattggtcatgccagcctgtgtccagtcctcttctgtcgtcaaaaat	1173	
Db	1018	GTTTTTCATTGCTAATGATTGGCATACCGCACTCTCGCCTGCTATCTAAGAGCCCTATTAC	1077	
Qy	1174	agaccaacagggtgtttacagagattcccgacagacccttgttatatacataaatttaacacat	1233	
Db	1078	CGGACAAATGTTTGAATGACGATATGCTCGCTCTGTGCTTGTGATACACAACAATTCCTCAT	1137	
Qy	1234	cagggtgtggagcctgcgaagtacatactcctgatctgggattgcctccctgaaatgggtatgga	1293	
Db	1138	CAGGTCGTGGCCCTGTAGACGACTTCCTCAATTTTGACTTTGCTTGGCTTGAACACTTAC-ATCGA	1196	

```

RESULT 12
US-09-196-390-5
: Sequence 5, Application US/09196390
: Patent No. 6307125
: GENERAL INFORMATION:
: APPLICANT: Block, Martina
: APPLICANT: Lorz, Horst
: APPLICANT: Lutticke, Stephanie

```

Qy	817	ccaagataacttaaatgggtctctctgataaaaaaactgcaagaagcgattatatacaactgcgaag	876
Db	1206	CCAAGGT-----ATGGGGACTATGAAGAAGCCTACGATGTCGGAGTC	1247
Qy	877	cacatlaagattccatgctttggggatcacatgaagtgcaccttttttcatagagtataga	936
Db	1248	CGAAATACTACAAGCGCTCGACAGGATATGGAAGTGAATTATTTCATGCTTATATC	1307
Qy	937	gacacagtcgagattgggtgttttcgatacatcgcgtcatatcacagaccaaggaagtatat	996
Db	1308	GATGGAGTGAATTTGTGTTCATTGACGCTCC-----TCCTCCGACACCGTCAGGAA	1361
Qy	997	ggagataaattttgggtcctttggtgataaatcagtttcagatcacacactcctttgtatgct	1056
Db	1362	GACATTTATGGGGCAGCAGACAGGAAATATGAAGCGCATGATTTTCTTCGAAGGCC	1421
Qy	1057	gcatgcgagccccactaatccttgaattgggagatatattatggaca---gaattgc	1113
Db	1422	GCTGTTGAGGTTCCATGGCACGTTCCATCGCGCGTTCCTTATTTGGGATGGAAATCTG	1481

Db	1482	GTGTTTATTGCAAAATGATTGGGACACGGGACATCCCTGGCCCTGCTATCTCTGAAAGCATATTATC	
Qy	1174	agaccatacagtggtttacagagattcccgcagcacccctgtttatataaatttagcacat	1233
Db	1542	AGGACCATGGTTTGATGCAGTACATCGTCGTCATTATGGTGATACATAAATCATCGCTCAC	1601
Qy	1234	cagggtgtggagcctcgaagtacatatcctgatactctggaatgacctcctggaatggtatgga	1293
Db	1602	CAGGGCCGTCGCCCTAGATGAATTCCTCGGTTTCACCGAGTTGCTCTGAGACACTACTCTGAA	1661
Qy	1294	gccttagaataagggtattctccagaaatgggccaaggagggcatgcccttgacaagggtgaggca	1353
Db	1662	CATTTC-----ACACTGTACGACCCCGTGGGTGGTGAACAC	1697
Qy	1354	gttaacttttgaagaggcagttgtgacagcagatcggattgtgacogtcaagtcagggt	1413
Db	1698	GCCAACTACTTCGCCCGCGCCTGAAGATGGCGGACCAAGTTGTCTGGTCGAGCCCGGG	1757

D5	1756	IACCTGTGGAGCGTGAAGACGGTGGAGGGCGGCTTGGGGCTTCACAGCATCATACAGCGG	181
Qy	1474	cgaaaagtgtattgaatgaattgataatggaattgacattaatgattggaaccccacc	1533
Db	1818	AACGACTGGAGAGCCCGGGCATCTGTCAAGGCATCGACACATGGATGGAGACCCGAG	1877
Qy	1534	acagacaaagtgtctc-----cctcatcattattctgtgatgacctctctgga----	1581
Db	1878	GTGGACGCCACCTCAAGTCGGACGGCTACACCACTTCTCCTGAGGACGCTGGACTCC	1937
Qy	1582	---aaggccaaatgaaagctgaattgcagaaggagtgggtttacctgtgaaggaggat	1638
Db	1938	GGCAAGCGGCAGTGCAAGGAGGCCCTGTCAGCGGAGCTGGGGCTTCAGGTCGGCGCCGAC	1997
Qy	1639	gttctctgattggcttatttgaagactggattaccagaaggcattgatctcattaaa	1698
Db	1998	GTACCGCTGTCTCGCTTATCGGCGCGCTGGACGGCAGACGGGCTGGAGATCATACCG	2057

DB 2058 GAGGCCATGCGCCCTGGATCTGTGAGCCAGGACGTGCAGCTGGTGATGCTGGGGCACCGGGGCG 2117

Qy	1819	tgggttggatttagttcttcaggtttcccccacagaataaactgcaggttgcgatatatattgta	1878
Db	2177		
Qy	2178	tggttgggggtttcccgcttggcccttggcgcaccgcgacatcagcgcggggcgagcgcgcgtctctc	2237
Db	2179		
Qy	1818	cacgaacttgagagacatgtctgcagcactttcgcagcgggagaccacgcacgaagtcgcgggg	2177
Db	2118		
Qy	1755	ccaaattttgaaggctcgatgatgatacccgaggttcaggtttacaaaggaataaatttcctggga	1818
Db	2119		

Search completed: March 28, 2002, 14:56:11
Job time: 9785 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2002, 13:48:26 ; Search time 3788.66 Seconds
(without alignments)
12818.690 Million cell updates/sec

Title: US-09-674-824-1
Perfect score: 2771
Sequence: 1 cgcactccactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues 34319436

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		Pending_Patents_NA_Main:*	
1:	/cgn2_6/ptodata/1/pna/US0611_COMB.seq.*	44:	/cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
2:	/cgn2_6/ptodata/1/pna/US06_COMB.seq.*	45:	/cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
3:	/cgn2_6/ptodata/1/pna/US07_COMB.seq.*	46:	/cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
4:	/cgn2_6/ptodata/1/pna/US080_COMB.seq.*	47:	/cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
5:	/cgn2_6/ptodata/1/pna/US081_COMB.seq.*	48:	/cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
6:	/cgn2_6/ptodata/1/pna/US082_COMB.seq.*	49:	/cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
7:	/cgn2_6/ptodata/1/pna/US083_COMB.seq.*	50:	/cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
8:	/cgn2_6/ptodata/1/pna/US084_COMB.seq.*	51:	/cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
9:	/cgn2_6/ptodata/1/pna/US085_COMB.seq.*	52:	/cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
10:	/cgn2_6/ptodata/1/pna/US086_COMB.seq.*	53:	/cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
11:	/cgn2_6/ptodata/1/pna/US087_COMB.seq.*	54:	/cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
12:	/cgn2_6/ptodata/1/pna/US088_COMB.seq.*	55:	/cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
13:	/cgn2_6/ptodata/1/pna/US089_COMB.seq.*	56:	/cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
14:	/cgn2_6/ptodata/1/pna/US090_COMB.seq.*	57:	/cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
15:	/cgn2_6/ptodata/1/pna/US091_COMB.seq.*	58:	/cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
16:	/cgn2_6/ptodata/1/pna/US092_COMB.seq.*	59:	/cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
17:	/cgn2_6/ptodata/1/pna/US093_COMB.seq.*	60:	/cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
18:	/cgn2_6/ptodata/1/pna/US094_COMB.seq.*	61:	/cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
19:	/cgn2_6/ptodata/1/pna/US095A_COMB.seq.*	62:	/cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
20:	/cgn2_6/ptodata/1/pna/US095B_COMB.seq.*	63:	/cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
21:	/cgn2_6/ptodata/1/pna/US095C_COMB.seq.*	64:	/cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
22:	/cgn2_6/ptodata/1/pna/US095D_COMB.seq.*	65:	/cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
23:	/cgn2_6/ptodata/1/pna/US096A_COMB.seq.*		
24:	/cgn2_6/ptodata/1/pna/US096B_COMB.seq.*		
25:	/cgn2_6/ptodata/1/pna/US096C_COMB.seq.*		
26:	/cgn2_6/ptodata/1/pna/US096D_COMB.seq.*		
27:	/cgn2_6/ptodata/1/pna/US096E_COMB.seq.*		
28:	/cgn2_6/ptodata/1/pna/US097A_COMB.seq.*		
29:	/cgn2_6/ptodata/1/pna/US097B_COMB.seq.*		
30:	/cgn2_6/ptodata/1/pna/US097C_COMB.seq.*		
31:	/cgn2_6/ptodata/1/pna/US098_COMB.seq.*		
32:	/cgn2_6/ptodata/1/pna/US099_COMB.seq.*		
33:	/cgn2_6/ptodata/1/pna/US6000_COMB.seq.*		
34:	/cgn2_6/ptodata/1/pna/US6001_COMB.seq.*		
35:	/cgn2_6/ptodata/1/pna/US6002_COMB.seq.*		
36:	/cgn2_6/ptodata/1/pna/US6003_COMB.seq.*		
37:	/cgn2_6/ptodata/1/pna/US6004_COMB.seq.*		
38:	/cgn2_6/ptodata/1/pna/US6005_COMB.seq.*		
39:	/cgn2_6/ptodata/1/pna/US6006_COMB.seq.*		
40:	/cgn2_6/ptodata/1/pna/US6007_COMB.seq.*		
41:	/cgn2_6/ptodata/1/pna/US6008_COMB.seq.*		
42:	/cgn2_6/ptodata/1/pna/US6009_COMB.seq.*		
43:	/cgn2_6/ptodata/1/pna/US6010_COMB.seq.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392.6	50.3	2491	17	US-09-345-214-5
2	1392.6	50.3	2491	29	US-09-743-980-5
3	1392.6	50.3	2491	42	US-09-094-436-9
4	1391.4	50.2	2008	17	US-09-345-214-12
5	1391.4	50.2	2008	29	US-09-743-980-12
6	1391.4	50.2	2008	42	US-09-094-436-10
7	1329.4	48.0	2606	25	US-09-654-617-385863
8	1329.4	48.0	2606	27	US-09-684-016-385863
9	1327.6	47.9	2392	14	US-09-077-564-1
10	1325.6	47.8	1749	18	US-09-402-254-52
11	1325.6	47.8	1752	24	US-09-625-406-12
12	1268	45.8	2216	25	US-09-654-617-451753
13	1268	45.8	2216	27	US-09-684-016-451753
14	1248.4	45.1	1620	24	US-09-625-406-20
15	1215.8	43.9	1528	17	US-09-345-214-6
16	1215.8	43.9	1528	29	US-09-743-980-6
17	1215.8	43.9	1528	42	US-09-094-436-10
18	1177.8	42.5	1854	64	US-60-312-544-3744
19	1015	36.6	1415	17	US-09-345-214-11
20	1015	36.6	1415	29	US-09-743-980-11
21	1015	36.6	1415	42	US-09-094-436-11
22	752.8	27.2	2360	23	US-09-606-304-9
23	747	27.0	1464	25	US-09-654-617-258645
24	747	27.0	1464	27	US-09-684-016-258645
25	712.2	25.7	1300	64	US-60-312-544-395
26	648.4	23.4	1758	23	US-09-606-304-3
27	614.6	22.2	1034	25	US-09-654-617-269739
28	614.6	22.2	1034	27	US-09-684-016-269739
29	555.6	20.1	619	58	US-60-253-654-32377
30	555.6	20.1	619	58	US-60-255-592-32377
31	547	19.7	595	58	US-60-253-654-5868
32	547	19.7	595	58	US-60-255-592-5868
33	540	19.5	660	53	US-60-209-830-55092
34	454.6	16.4	558	23	US-09-619-643-16489
35	438.2	15.8	555	23	US-09-619-643-16488
36	436.2	15.7	489	58	US-60-253-654-32385
37	436.2	15.7	489	58	US-60-255-592-32385
38	387.8	14.0	430	26	US-09-667-188A-1103
39	387.8	14.0	430	48	US-60-155-008-1836
40	353.6	12.8	469	53	US-60-207-458-39367
41	349	12.6	2478	14	US-09-077-564-3

42 344.8 12.4 518 53 US-60-207-458-64619 Sequence 64619, A
c 43 338 12.2 539 46 US-60-135-951-2922 Sequence 2922, Ap
44 333.2 12.0 529 17 US-09-371-146A-83476 Sequence 83476, A
45 311.6 11.2 334 58 US-60-253-654-5901 Sequence 5901, Ap

ALIGNMENTS

RESULT 1

US-09-345-214-5
; Sequence 5, Application US/09345214
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; FILE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Zea mays
US-09-345-214-5

Query Match 50.3%; Score 1392.6; DB 17; Length 2491;
Best Local Similarity 81.4%; Pred. No. 6e-119;
Matches 1677; Conservative 0; Mismatches 369; Indels 13; Gaps 5;
Qy 166 cagcaccgagctggagagagagcttcgcccgcccgccagcagcgcgggcgatccacc 225
Db 22 caccgtcgtagaagacacgggagcagcaccgccgagctcgctcgctccctca 81
Qy 226 gtccgtcgctcgcaactctccgctctctccctctccctctgctcccgccgagcgcg 285
Db 82 ctctctcccgccgc-gatccacggcccgcccgcccgccgctctgctgctccctc 140
Qy 286 gcgacggcgctcgccgctcgctcgcccgcccgcccgcccgcccgcccgcccgccg 345
Db 141 cgcaatggcgagccctcg--ccgtggcgcgctgctctctccctcgcgcgccgc 198
Qy 346 acggcgcccgccgctcgctcgctcgctcgctcgctcgctcgctcgctcgctcgctc 405
Db 199 tggccggcccgctcgccgagccggcgcccgcccgcccgcccgcccgcccgccgc 258
Qy 406 cgctacgctcgccgagctcagcagggggcccgcccgcccgcccgcccgcccgccgagc 465
Db 259 cggtcgctcgcgagctgagcagggggggggggggggggggggggggggggggggg 318
Qy 466 ctggcccgcccgctcgctcgagctctctcgcccgcccgcccgcccgcccgcccgccg 525
Db 319 ctggcgcccgcccgctcgctcgcccgctctctcg---cgcccgcccgcccgcccgccg 375
Qy 526 ccggcccgcccgagcagcccgcccgagcccgcccgcccgcccgcccgcccgcccgcc 585
Db 376 ccggatcgagcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcc 431
Qy 586 ctgctcgaaaggggagctgagagcttcctcgacagcagataatgtggtcgcaagtgcag 645
Db 432 --acctgaagggattgctgaaggttcctcgataacacagatgtgtggcaagtgcagaa 489
Qy 646 gattctgagatcatgatcgatgagcaactcaagctaaagtctacagtcgcatctg 705
Db 490 gattctgagatctggttggaagagagcagagctcgagctaaagttaacaagaagctgtc 549
Qy 706 ttgtgactggtgaagctgctctctatgcaaaagtcagggggttgaggagatgttgtgt 765
Db 1630 acagagtcgactctcaaggataaaatttcgtggatgggttggtatttagtgttccagttcc 1699

Db 550 ttgttaaccggcggaagcttctccttatgtcaaaagctctgggggctctaggagatgttttggtt 609
Qy 766 tegtaccaaatgtctctgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 825
Db 610 tcaatggcagttgtctgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 669
Qy 826 ttaaatgggtctctctgtataaaactatgtcaaaagcattatcacactgcgagcagcattcaag 885
Db 670 ttaaatgggtacctccgataaagaaattatgcaaatgcattttcacagaaaaacacattcgg 729
Qy 886 attccatgctttgggggacatcacatgaagtgtacccctttttcatgtatgataagacaacgtc 945
Db 730 attccatgctttggcggtgaacatgaagtgtacccctttttcatgtatgataagatcagtt 789
Qy 946 gattgggtgtttgtcgatcatccgtcatatcacagaccaggaagtgtttatatggagataat 1005
Db 790 gactgggtgtttgtgtatcatccctcatatcacagacctggaaatttatatggagataag 849
Qy 1006 ttgtgtctttgtgtataatcagttcagatcacacactcctttgctatgctgcatcgag 1065
Db 850 ttgtgtctttgtgtataatcagttcagatcacacactcctttgctatgctgcatgtgag 909
Qy 1066 gcccaactatccttggaattgggagatatattatggacagaattgcattgtttgtgtg 1125
Db 910 gctccttgatccttgaattgggagatatattatggacagaattgcattgtttgtgtg 969
Qy 1126 aagcattgtgcatgccagcctgtgtccagtccttctgtcgtcaaaatatagacatacgtt 1185
Db 970 aatgattggcatgccagctcagtgccagtccttctgtcgtcaaaatatagacatacgtt 1029
Qy 1186 gtttacagagattcccgagcagcaccctgtttatataaatttagcacatcaggtgtggag 1245
Db 1030 gttataaagactcccgagcagcattctgttaatacaataatttagcacatcaggtgtggag 1089
Qy 1246 cctgcaagtacatcctcgtggtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1305
Db 1090 cctgcaagcacatcctcgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1149
Qy 1306 gtatttcccaagattggcagggagcagcctcctgtcaaaaggtgtgtgtgtgtgtgtgtgt 1365
Db 1150 gtatttccctgaattggcgaggagcagcctcctgtcaaaaggtgtgtgtgtgtgtgtgtgt 1209
Qy 1366 aaaggcgagttgtgacagcagatcggattgtgacctgacgtcagtgaggtttatcagtgga 1425
Db 1210 aaaggcgagttgtgacagcagatcggaaatcgagctgtcagtaagggtttatcagtgga 1269
Qy 1426 gtcaacaactgtgaagtgagcagggcctcaatgagctcttaagctcccgcaaaagtgtga 1485
Db 1270 gtcaacaactgtgaagtgagcagggcctcaatgagctcttaagctcccgcaaaagtgtga 1329
Qy 1486 ttgaattgggaattgtaaatgggaattgacattaatgttggaaccccccacacagacaagtgt 1545
Db 1330 ttgaacgggaattgtaaatgggaattgacattaatgttggaaccccccacacagacaagtgt 1389
Qy 1546 ctccctcattcatttctgtgagtaactctctggaagggccaaatgttaaagctgaaattg 1605
Db 1390 atccctcgtcatttctgtgtgagtaactctctggaagggccaaatgttaaagctgaaattg 1449
Qy 1606 cagaagggagttgggtttacctgtaaggagagatgttctctgtgattggtcttatttggaaga 1665
Db 1450 cagaagggagctgggtttacctgtaaggagagatgttctctgtgattggtcttatttggaagg 1509
Qy 1666 ctggattaccgagaagagcattgatctcattaaattgcccattccagagctcatgagggag 1725
Db 1510 ttggattatccagaagggcattgatctcattaaattgcccattccagagctcatgagggag 1569
Qy 1726 gacgtgcaattgtcattgctgtggatcctgggagatccaaatttttgaggcctggatgagatct 1795
Db 1570 gatgtcaattgtcattgctgtggatcctgggagatcctgggagcctggagattggatgagatct 1629
Qy 1786 accgagtcagttacaaggaataattccgtggaatgggtgtgatttagtgttccagtttcc 1845
Db 1630 acagagtcagttccaaggataaaatttcgtggatgggtgtgatttagtgttccagtttcc 1699


```
Qy 1246 cctgcaagtacatatcctgactggtggttgcctcctcaatggtatgagcctttagaatgg 1305
Db 1090 cctgcaagcacatcctgactggttgcctcctcaatggtatgagcctttagaatgg 1149
Qy 1306 gatttccagaatgggcaaggagcattgccttgacaaagggtgaggcagtttaattttg 1365
Db 1150 gatttccagaatgggcaaggagcattgccttgacaaagggtgaggcagtttaattttg 1209
Qy 1366 aaaggagcagtttgacacagacatcggtattgtgacctcagtcagggtttattcaatggag 1425
Db 1210 aaagggtcagtttgacacagacatcggtattgtgacctcagtcagggtttattcaatggag 1269
Qy 1426 gtcaaaactgtgaagtgacagggcctcaatgagctcttaagctcccaaaaaagtgtta 1485
Db 1270 gtcaaaactgtgaagtgacagggcctcaatgagctcttaagctcccaaaaaagtgtta 1329
Qy 1486 ttgaatggaattgtaaatgaattgacattaatgattggaaacccaccacagacaagtgt 1545
Db 1330 ttaaacggaattgtaaatgaattgacattaatgattggaaacccaccacagacaagtgt 1389
Qy 1546 ctccctcatcatattctgtcgatgacctctctggaagcccaaatgttaagctgaattg 1605
Db 1390 atccccgtcatattctgtgtgacacctctctggaagcccaaatgttaagctgaattg 1449
Qy 1606 cagaaggagttggtttacctgtaaaggaggtgttctctgtattggtctttattgaaaga 1665
Db 1450 cagaaggagttggtttacctaataaggcctgatttctctgtattggtctttattgaaaga 1509
Qy 1666 ctgattaccagaaaagcattgctatcattaaatggcccaattccagagctcatgaggag 1725
Db 1510 ttggattacagaagcattgctatcattaaatggcccaattccagagctcatgaggag 1569
Qy 1726 gacgtgcaattgtcattgctgctgattctgggacccaattttgaaagctggaatgatct 1785
Db 1570 gatttcaattgtcattgctgctgattctgggacccaattttgaaagctggaatgatct 1629
Qy 1786 accgagtcagtttacaaggataaattccgtggtggttggattttagtttccagtttcc 1845
Db 1630 acagagtcagtttacaaggataaattccgtggtggttggattttagtttccagtttcc 1689
Qy 1846 cacagaataactcaggttgcgatatattgttaatgccatcagatttgaaccttgcgt 1905
Db 1690 caccgataactcgcgcgcgatatattgttaatgccatcagatttgaaccttgcgt 1749
Qy 1906 cttaatcagctatatgctatgcaaatggttacagttcctgtgatttcaatggaactggggc 1965
Db 1750 ctcaatcagctatatgctatgcaaatggttacagttcctgttgcctatgcaactggggc 1809
Qy 1966 ctccgagacacagtcagaccttcaaccttttggtgcaaaaaggagaggggtacaggg 2025
Db 1810 cttagagatccgtggagaaacttcaaccttttgcgttgagaaaggagaggggtacaggg 1869
Qy 2026 tgggcttctcaccgttaaccgttgacagaatgttgtggaacttgcgaaccgcgatgtcg 2085
Db 1870 tgggcaatcgaccctcaccacagaaaaacatgttgtggacattgcgaactgcaatatct 1929
Qy 2086 acattcaggggacacaagcctcctgggaggggctcatgaagcggaggtacgaaagac 2145
Db 1930 acatacaggggacacaagcctcctgggagggcctaataagcagggatgttcaaaaagac 1989
Qy 2146 catacgtgggacattgcc-ccgagcagttacgagcagatattctgagttggccttcgtgac 2204
Db 1990 ttcacgtgggacattgccgcgtgaaacaatacagcaaaattcttccagtgggaccttcgat 2049
Qy 2205 caacctcagctcatgtaga 2223
Db 2050 cgacctatgtcatgtaaa 2068
```

RESULT 4

US-09-345-214-12

; Sequence 12, Application US/09345214

; GENERAL INFORMATION:

```
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-345-214-12
```

```
Query Match 50.2%; Score 1391.4; DB 17; Length 2008;
Best Local Similarity 83.8%; Pred. No. 8e-119;
Matches 1612; Conservative 0; Mismatches 301; Indels 10; Gaps 3;
```

```
Qy 302 ccgggtgctcgtcccccagcgtccgctcgtcgcgccgatccggcgacggcgccggcggt 361
Db 12 cctcgccctggcgccgctgctcctcctcgtcgcgccgatccggcgacggcgccggcggt 71
Qy 362 ccgctcgtcgtccgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 421
Db 72 ggcacggcgcgccgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 131
Qy 422 tcagcaggagagggccgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 481
Db 132 tgagcaggagagggccgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 191
Qy 482 tgccaggttctcgtcgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 541
Db 192 tgcccggttctcgtcgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 248
Qy 542 cgccctcgtcgcgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 601
Db 249 cgccctcgtcgcgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 302
Qy 602 ctgagatttccatcagcagcataattgttcgttcgttcgttcgttcgttcgttcgttcgttc 661
Db 303 ctgagatttccatcagcagcataattgttcgttcgttcgttcgttcgttcgttcgttcgttc 362
Qy 662 atcgaaatgagcaacctcaagctaaagtacacgtagcatcgttcgttcgttcgttcgttcgttc 721
Db 363 ttggaaaggagcaagctcagcgttaagtaaacacaaagcatgttcgttcgttcgttcgttcgttc 422
Qy 722 ctgctccttatgcaagtcaggggttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 781
Db 423 ctctccttatgcaagtcaggggttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 482
Qy 782 ttgctcgttcgttcagcaggttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 841
Db 483 ttgctcgttcgttcagcaggttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 542
Qy 842 ataaaaactatgcaagtcattatcacctgcgaagacacacattaaagattccattcgttcgttc 901
Db 543 ataaagaattatgcaaaatgcatctttacacagaaaaacacattcggattccattcgttcgttc 602
Qy 902 gatcacatgaagtgcaccttttttccatgatgatagacacacgttcgttcgttcgttcgttcgttc 961
Db 603 gtgaacatgaagtgcaccttttttccatgatgatagacacacgttcgttcgttcgttcgttcgttc 662
Qy 962 atcatcgttcatacacagacaggaagtttatatgatgagataattttgttcgttcgttcgttcgttc 1021
Db 663 atcatcctcctatacacagacgtgaaatttatatgatgagataattttgttcgttcgttcgttcgttc 722
Qy 1022 ataatcagttcagatacacactcctttgtctatgctgcgtgcagggccccacataccttcgttc 1081
```



```

      832 GCAAGCACATCTGACCTGGTGGCCACCTGAATGGTATGAGCTCTGGAGTGGGTA 891
      1309 ttctcagaatggcagagagcgcaccccttgacaagggtaggcagcttaacttttgaaa 1368
      892 TTCCCTGAATGGGAGAGGATGCCCTTGACAAGGGTGAGGCAGTAAATTTTTGAAA 951
      1369 ggaagcagttgtgacagcagatcgagattgtgacgctcaagtgcaggttattcatggaggtc 1428
      952 GGTGCAGTTGTGACAGCAGATCGAATCGTCACTGCTAGTAAGGGTATTCCTGGGAGGTC 1011
      1429 acaactgctgaaggtggacagggcctcaatgagctcttaagctctcccgaaaaaagtgtattg 1488
      1012 ACAACTGCTGAAGGTGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAAAGAGTGATTA 1071
      1489 aatggaattgtaaatgaaatgacattgaattgaattgaacccaccacacagaagtgtctc 1548
      1072 AACGGAATGTAAATGGAATGACATTAATGATTGGAAAGCCCTGCCACAGACAAATGTATC 1131
      1549 cctcatcattattctgtcgatgacctctctggaaaggccaaaatgtaagctgaattgcag 1608
      1132 CCCGTCTATATCTGTGTGATGACCTCTCTGGAAGGCCAAATGTAAAGCTGCATTGCAG 1191
      1609 aaggagttgggtttacctgtcaaggagagatgttctctctgattggtcttattggaagactg 1668
      1192 AAGGAGCTGGGTTTACCTATTAAGGCCCTGATGTTCTCTGATTGGCTTTATTGGAAGGTTG 1251
      1669 gattaccagaaagcattgacctcattcaataaagtccattccagagctcatgagggagagac 1728
      1252 GATTATCAGAAAGGCATGTGATCTCATTAACCTTATCATACAGATCTCATGCGGGAAGAT 1311
      1729 gtgcaattgtcatgctgtgacatgggagatccaaattttgaaagctggatgagatctacc 1788
      1312 GTTCAATTTGTCATGCTTTGGATCTGGTGGACCCAGAGCTTGAAGATTGGATGAGATCTACA 1371
      1789 gagtcgagttacaaggataaattccgtggatgggttgagatttagtgcagtttccacc 1848
      1372 GAGTCGATCTTCAAGGATAAATTTGCTGGATGGGTGGATTGATTTAGTGTCCAGTTTCCCAC 1431
      1849 agaataactgcaggttgcgatataattgttaatgcatcgagatttgaaccttgcggtctt 1908
      1432 CGAATAACTGCGGCTGGCATATATTGTTAATGCCATCCAGATTGCAACCTTGTGGTCTC 1491
      1909 aatcagctatatgtatgcaataatggtacagttctctgttagtctcatgaaactgggggctc 1968
      1492 AATCAGCTATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1551
      1969 cgagacagtcgagacctcaacctttgtgcaaaagagagagaggttacaggtgg 2028
      1552 AGAGATACCGTGGAGAACTTCAACCCCTTCGGTGAGATGGAGAGCAGGGTACAGGGTGG 1611
      2029 gcgttctcacccgctaaccctgtggacaagatg-ttgtgggcatctgcgaaccgcgagtgcgac 2087
      1612 GCATTCGACCCCTTAACACAGAAAACATGTTGTGGACATTCGAACTGCAATATCTAC 1671
      2088 attcaggagacacaagcgtctctggggggggtctcatgaagcgaggtcatgacgaagacca 2147
      1672 ATACAGGGAACAAAGTCTCTCTGGGAAGGGCTAATGAAGCGAGGCATGTCAAAAGACTT 1731
      2148 tacgtggagaccatgccc 2165
      1732 CACGTGGGACCATGCCGC 1749
  
```

RESULT 12

```

US-09-654-617-451753
; Sequence 451753, Application: US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
  
```

```

; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 451753
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Sorghum bicolor
US-09-654-617-451753
  
```

```

Query Match      45.8%; Score 1268; DB 25; Length 2216;
Best Local Similarity 86.5%; Pred. No. 1.5e-107;
Matches 1422; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

QY 598 attgctgaggtatcccatcgacagcataaattgtgtgctgcgaagtgcagcagattctgagatc 657
Db 1 atgctggaaggttcccatcgatgagacagtagtctgtgccaagtgcgaagattctgagatc 60
QY 658 atggtatgcgaatgagcaaacctcaagctaaagttaacacgtacgctgctgtgttgactggtt 717
Db 61 gtggttggaagagcaagctcagctaaagttaacacagcattctctctgcaactggc 120
QY 718 gaagctgctctctatgcaaaagtcagggggttgggagatgtttgtgtgttctgttaccact 777
Db 121 gaagctctctctatgcaaaagtcagggggttctagggatgtttgtgtgttctatgccagtt 180
QY 778 gctctgtgctgctgtgtccaccagtgatgtgtgttaatgccaagatacttaaatgggtcc 837
Db 181 gctctgtgctgctgtgtccaccgtgtgatgtgttaatgccagataatttaaatgggtacc 240
QY 838 tctgataaaaactatgcaaaagcattatcacatgcagcagacacattaaagattccatgcttt 897
Db 241 tctgataaagattcacgcaaatgcatcttcaacagaaaaagcacattccgattccatgcttt 300
QY 898 gggggatcacatgaagtgcacctttttccatgagtagatagacacaacgtcgattgggtgttt 957
Db 301 ggggtgaacatgaagttaacctttttccatgagtagacagagatcagttgactgggtgttt 360
QY 958 gtgcatactccgtcattatccacagaccaggagtttataatggagataattttgggtgtttt 1017
Db 361 gttgatactccctcattatccacagaccctggaaatttataatggagataattttgggtgttt 420
QY 1018 ggtgataatcagtttcagatcacacctcttctgctatgctgcatgctgcagagcccaactaatc 1077
Db 421 ggtgataatcagtttcagatcacacctcttctgctatgctgcatgctgcagctcctttggtc 480
QY 1078 cttgaaattggagagatataattatggacagaattgcattgtgttgcaacgattggcat 1137
Db 481 cttgaaattggagagatataattatggacagaattgcattgtgttgcaacgattggcat 540
QY 1138 gcaagcctgtgacagctcctcttctgctgcaaaatatagaccatacagctgtttacagagat 1197
Db 541 gcaagcctgtgacagctcctcttctgctgcaaaatatagaccatacagctgtttacagagat 600
QY 1198 tcccgacagcctctgttatataataatttagcacatcagggtgtggagcctgcaagtaca 1257
Db 601 tcccgacagcctctgttatataataatttagcacatcagggtgtggagcctgcaagtaca 660
QY 1258 tatcctgatctgggattgcctcctgaaatggatggagcttttagaatgggttatttccagaa 1317
Db 661 tatcctgatcctgggttgcacacctgaaatggatggagctctggagtggttattccctgaa 720
QY 1318 tgggcaaggagagcattgaccttgacaaggggtgagcagtttaacttttgaaaggagcagtt 1377
Db 721 tgggcaaggagagcattgaccttgacaaggggtgagcagtttaacttttgaaagggtgcagtt 780
QY 1378 gtgacagcagatcgattgtgacctgcagtcaggggttattcatggagggttcacaaactgct 1437
Db 781 gtgacagcagatcgaaattgtgactgtcagtaagggttattcatggagggttcacaaactgct 840
QY 1438 gaaggtggacagggcctcaatgagctcttaagctccccgaaaaaagctattgaaatggaaatt 1497
Db 841 gaaggtggacaggggtctcaatgagctcttaagctccccgaaaaaagctattgaaatggaaatt 900
  
```


QY 1498 gtaaatggaattgacattaatgattggaacccaccacagacacagtgctccctcatcat 1557
Db 901 gtaaatggaattgacattaatgattggaacccctgcagcgacaaaatgtatcccttgcat 960
QY 1558 tattctgtcagatgacctctctctggaaggccaaatgtaaaagctgaattgcgaaggagtg 1617
Db 961 tattctgtgatgacctctctctggaaggccaaatgtaaaagctgattgcgaaggagctg 1020
QY 1618 ggtttacctgaaggaggagtgctctctctgattggtttatttattggaagactggattaccag 1677
Db 1021 ggtttacctgaaggaggagtgctctctctgattggtttatttattggaagactggattaccag 1080
QY 1678 aaaggcattgatcatataaaatggccattccagagctcatgaggagcgagctgcaattt 1737
Db 1081 aaaggcattgatcatataaaatggccattccagagctcatgaggagcgagctgcaattt 1140
QY 1738 gtcagcttgatctggggatccaaatttttgaaggctggaagactaccggagctgagtt 1797
Db 1141 gtcagcttgatctggggatccaaatttttgaaggctggaagactaccggagctgagtt 1200
QY 1798 tacaaggataaaatcccgtagtggtgatttagttccagtttccacagaaataact 1857
Db 1201 tacaaggataaaatcccgtagtggtgatttagttccagtttccacagaaataact 1260
QY 1858 gcaggttgcgatatattgttaatgccatcgagatttgaaccttgcggtcttaacagcta 1917
Db 1261 gcaggttgcgatatattgttaatgccatcgagatttgaaccttgcggtcttaacagcta 1320
QY 1918 tatctatgcaatagtgacagttctctgtagttcatggaactgggggctccgagacaca 1977
Db 1321 tatctatgcaatagtgacagttctctgtagttcatggaactgggggctccgagacaca 1380
QY 1978 gtccagaccttcaaccttttgggtcgaaggagaggggtacaggttggcgcttctca 2037
Db 1381 gtccagaccttcaaccttttgggtcgaaggagaggggtacaggttggcgcttctca 1440
QY 2038 ccgctaaccgtgacaaagatg-ttgtgggcattgcgaacgcgagatcgcacattcaggga 2096
Db 1441 ccgctaaccgtgacaaagatg-ttgtgggcattgcgaacgcgagatcgcacattcaggga 1500
QY 2097 gcaaacgcgtctctggaggggctcatgaagcagagcatgacgaaagaccatcagtgga 2156
Db 1501 gcaaacatctctctggaggggctcatgaagcagagcatgacgaaagaccatcagtgga 1560
QY 2157 ccagtcg-ccgagcagtagcagagatcttcagtggtggcctctctgtagcacaacctcagct 2215
Db 1561 ccagtcgctgaacatacgaacaaatctccagtggtggccttcatcgatcgaccctatgt 1620
QY 2216 catgtagacgggagctggggaggt 2239
Db 1621 catgtaaaggggaccacaagtgt 1644

RESULT 13

US-09-684-016-451753
; Sequence 451753, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 451753
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Sorghum bicolor
US-09-684-016-451753

Query Match 45.8%; Score 1268; DB 27; Length 2216;
Best Local Similarity 86.5%; Pred. NO. 1.5e-107;
Matches 1422; Conservative 0; Mismatches 220; Indels 2; Gaps 2;
QY 598 attgtgagagattcccatcgacacataattgtgctgcaagtgagcagagattctgagatc 657
Db 1 attgtgagagattcccatcgatgagacagtagttgtggaagtgagcaagattctgagatc 60
QY 658 atggatgcgaatgagcaaacctcaagctaaagtacacgtagcatcgtgtttgtgactggt 717
Db 61 gtggttggaaaggagcaagctcgagctaaagtaacacaaagcattgtctttgtaactggc 120
QY 718 gaagctgctccttatcaaaagtcagggggttggagatgtttgtgttcgtttacaacatt 777
Db 121 gaagctgctccttatcaaaagtcggtggtctagggagattgtgtgttcattgccagtt 180
QY 778 gctcttgcgtcgtggtcacccagtgatggttaatgccaaagatacttaaaatgggtcc 837
Db 181 gctcttgcgtcgtggtcacccgtgtaggttgaatgccagataatttaaatggtacc 240
QY 838 tctgataaaaaactatgcgaaggcattatatacactgcgaagcacacattaaattccatt 897
Db 241 tctgataaaatcagcaaatgcattttacacagaaagcacattcgtatccattcatt 300
QY 898 gggggatcacatgaagtgcacctttttcatgagtagagacacacacgtcagtggtgttt 957
Db 301 ggcgggtgaacatgaagtgcacctttttccatgagtagacagagattcagttgacgggttt 360
QY 958 gtccatcatccgtcatatcacagaccaggaagtttatatggagataatttttgggtgtttt 1017
Db 361 gttgatcatccctcatatcacagacctggaaatttatatggagataagtttgggtgtttt 420
QY 1018 ggtgataaatcagttcagatatacacactccctttgctatgctgcagcgagcccaacta 1077
Db 421 ggtgataaatcagttcagatatacacactccctttgctatgctgcagcgctcctttggct 480
QY 1078 ctgtgaattgggagagatatattatgacagaaatttgatgtgtgtgacattggcat 1137
Db 481 ctgtgaattgggagagatatattatgagacagaatttgatgtgtgtgagatgagtcac 540
QY 1138 gccagccttgcctcagtccttcttgcgcacaaatatagaccatacgggtgtttacagagat 1197
Db 541 gccagtcctagtccagtccttcttgcgcacaaatatagaccatacgggtgtttataaagac 600
QY 1198 tccgcagcacacctgtttatataataatttagcacatcagggtgtgagagcctgcaagtaca 1257
Db 601 tccgcagcatctctgttaataataatttagcacatcagggtgtgagagcctgcaagcaca 660
QY 1258 tatctctgattgggattgctcctcgaattggtatgagccttagaatgggtatttccagaa 1317
Db 661 tatctctgattgggattgctcctcgaattggtatgagcctggaatgggtatttccctgaa 720
QY 1318 tgggcaaggagagcagtccttgcacaaagggtgaggcagtttaactttttgaaaaggacgtt 1377
Db 721 tgggcaaggagagcagtccttgcacaaagggtgaggcagtttaactttttgaaaaggacgtt 780
QY 1378 gtgacagcagatcgattgtgacctcagtcaggtatttcatgagagagtcacaactgt 1437
Db 781 gtgacagcagatcgattgtgacctcagtcaggtatttcatgagagagtcacaactgt 840
QY 1438 gaaggtgacagggcctcaatgagctcttaagctcccgaaaaagtgtattgaatggaatt 1497
Db 841 gaaggtgacagggcctcaatgagctcttaagctcccgaaaaagtgtattgaatggaatt 900
QY 1498 gtaaatggaattgacataatgatttggaaacccacacagacagaagtgtctcctccatcat 1557
Db 901 gtaaatggaattgacataatgatttggaaacccctgcagcggaacaaatgtatcctctgtcat 960
QY 1558 tattctgtcagtagcctctctggaaggcgcaaatgtaaagctgaatttcagagagagttg 1617
Db 961 tattctgttagtagcctctctggaaggcgcaaatgtcaaaagtcgatttcagagagagctg 1020
QY 1618 ggttacctgtaaggaggagtagttcctctgattggtcttatttgggaagcaggtattaccag 1677

Db 1021 gggttaacctaaaggcctgaagtctcctctgattggtttatttggaagattggattacag 1080
Qy 1678 aaaggcattgctcattaaatggccattccagagctcatgagggagcgctgcaattt 1737
Db 1081 aaaggcattgctcattaaatggccattccagagctcatgagggagcgctgcaattt 1140
Qy 1738 gtcatgcttgatctggggtaccaaattttgaaggctggatgagatctaccgagtcgagt 1797
Db 1141 gtcatgcttgatctggggtaccaaattttgaaggctggatgagatctaccgagtcgagc 1200
Qy 1798 tacaagataaaattccgctgagtgagggttgatttagttccagtttccacacagaataact 1857
Db 1201 tcaagataaaattccgctgagtgagggttgatttagttccagtttccacacagaataact 1260
Qy 1858 gcaggttgcgatatattgttaagccatcgagatttgaaccttgcggtcttaataatcagcta 1917
Db 1261 gcggctgcgatatattgttgaagccatcgagatttgaaccttgcggtcttaataatcagcta 1320
Qy 1918 tatgctatgaatatggtacagttctctgtagtctgaactgggggacctcccgagacaca 1977
Db 1321 tatgctatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1380
Qy 1978 gtgagaccttcaaccttttgggtgcaaaaggagagagagagagagagagagagagagagag 2037
Db 1381 gtgagaccttcaaccttttgggtgcaaaaggagagagagagagagagagagagagagagag 1440
Qy 2038 ccgtaaccgtggacaagatg-ttggtggcgattgcaacccgagtgccgacattcagggga 2096
Db 1441 cccctaaccaggaacacatgttgggtgagacattgcaacccgagtgccgacattcagggga 1500
Qy 2097 gcaaaagccgtctctggaggggctcatgaagcgagggcatgacgaaagacacatacggtggga 2156
Db 1501 gcaaaatctctctggaggggctcatgaagcgaggggcatgacgaaagacacatacggtggga 1560
Qy 2157 ccattgcc-cgagagcagtgacgagacatcttcgagtgagggccttcgtagacacacccctacgt 2215
Db 1561 ccattgccctgagac 1620
Qy 2216 catgtagacggggactgggaggt 2239
Db 1621 catgtaagggggacacacagtggt 1644

RESULT 14

US-09-625-406-20
Sequence 20, Application US/09625406
GENERAL INFORMATION:
APPLICANT: Keeling, Peter
APPLICANT: Guan, Hanping
TITLE OF INVENTION: Starch Encapsulation
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/625,406
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/941,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P

REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-09-625-406-20

Query Match 45.1%; Score 1248.4; DB 24; Length 1620;
Best Local Similarity 86.2%; Pred. No. 9.8e-106;
Matches 1392; Conservative 0; Mismatches 221; Indels 1; Gaps 1;
Qy 553 gacgcggcgctgggggaactcgcgcgacacctctgctgaaggattgctgaggattcc 612
Db 4 gtcgcggagctgagcaggagacacctcggtctcgaaacctgaaggattgctgaggattcc 63
Qy 613 atcgacgacgataatgtggctgcaagtgcagtgagcaggtctgagatcatggtgcgaatgag 672
Db 64 atcgatnaacacagtagttgtggcaagtgacgaagattctgagattgtggttggaaggag 123
Qy 673 caacctcaagctaaagtacacgtagcatcggtgttggactggtgagagctgctccttat 732
Db 124 caagctcgagctaaagtgtaacacaaagcattgtcttttaacccggcgaagcttctccttat 183
Qy 733 gcaaaagtcagggggtggagatgttgggttcgttaccgaattgctctgctgcgtcgt 792
Db 184 gcaaaagtcagggggtggagatgttgggttcgttaccgaattgctctgctgcgtcgt 243
Qy 793 ggtcaccgagtgatggtgttaagcgaagatacttaaatgggtcctctgataaaactat 852
Db 244 ggtcaccggtgatggtgttaagcgaagatacttaaatgggtcctctgataaaactat 303
Qy 853 gcaaaagcattatcacctgcgaagcagacattcaagattccattgggggagacacatgaa 912
Db 304 gcaaatgcattttacacagaaacacacattcggaattccattggcggtgaacatgaa 363
Qy 913 gtgacctttttcatgagtagagacacagtcgagtggtgttgcgatcatccgtca 972
Db 364 gttacctttctccatgagtagatagagattcagttgactgggtgttgggtgattcattccatca 423
Qy 973 tatcacagaccaggaagtattataggagataattttgggtccttttgggtgataatcagttc 1032
Db 424 tatcacagacctggaaattttataggagataattttgggtccttttgggtgataatcagttc 483
Qy 1033 agatacacactccttgcgtatgctgcgagggcccccactaatccttgatgggagga 1092
Db 484 agatacacactccttgcgtatgctgcgagggcccccactaatccttgatgggagga 543
Qy 1093 tatatttatggacagaattgcatgtttgtgaacagattggcagccttggtgcca 1152
Db 544 tatatttatggacagaattgcatgtttgtgaacagattggcagccttggtgcca 603
Qy 1153 gtccttctgctgcaaaatagaccatacgggtgttttaccagagattccccagacacccctt 1212
Db 604 gtccttctgctgcaaaatagaccatacgggtgttttaccagagattccccagacacccctt 663
Qy 1213 gttatacataatttagcacatcaggggtgtgagcctcagatcatcctgactgctggga 1272
Db 664 gtaatacataatttagcacatcaggggtgtgagcctcagatcatcctgactgctggga 723
Qy 1273 ttgcctcctgaaatgtagtgagccttttagaatgggtatttccagaaatgggcaagagagcat 1332

```

Db 724 TTGCCACCTGAATGTGATGGAGCTCTGGAGTGGGTATTCCCTGAATGGCGGAGGAGCAT 783
QY 1333 gcccttgacaaaggtgagcagtttaactttttgaaagagcagttgtgacagcagatcgg 1392
Db 784 GCCCTTGACAAGGGTGAGGAGCAGTTAATTTTGAAGGTGCAGTGTGTGACAGCAGATCGA 843
QY 1393 attgtgaccgtcagtcaggttattcatggagggtcacactgctgaaaggtggacagggc 1452
Db 844 ATCGTGACTGTCAGTAAGGGTTATTTCGTGGAGGTCACAACCTGCTGAAGGTGGACAGGC 903
QY 1453 ctoaatgagctcttaagctcccgaaagagtgattgaaatggaattgaaatgac 1512
Db 904 CTCAATGAGCTCTTAAGCTCCAGAAAGAGTATTAAACGGAATTGTAATGGAATTGAC 963
QY 1513 attaatgattgaaaccccccacaaagagtgatctccctcatcattattctctcagatgc 1572
Db 964 ATTAATGATTGAAACCTCCACAGACAATGATCCCCGTCAATATTCTGTGTGATGAC 1023
QY 1573 ctctctggaagggccaaatgtaagctgaattgcagaagaggttggttttacctgtaaag 1632
Db 1024 CTCTCTGGAAGGCCCCAAATGTAAGGTGCATTGCAGAGGAGCTGGGTTTACCTATAAG 1083
QY 1633 gaggatgtctctctgattgcttattggaagactggattaccagaaagcattgactc 1692
Db 1084 CCTGATGTTCTCTGATTGGCTTTATTGGAAGGTTGGATTATCAGAAAGGCATTGATCTC 1143
QY 1693 attaaatggccattccagactcatgaggagagtgacgagtgcaattgtcattgtgactc 1752
Db 1144 ATTCAACTTATATACCAATCTCATGCGGGAAGATGTTCAATTTGCTGATCTGTGATCT 1203
QY 1753 gggatcccaattttgaaaggtggatgagatctaccgagctcaggttcacaaagataaattc 1812
Db 1204 GGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAATT 1263
QY 1813 cgtgagatgggttgattgattcagttccagtttccacagaaataactcagagttgcgata 1872
Db 1264 CGTGGATGGGTGGATTTAGTGTTCAGTTTCCACCGAATAAATACCTCCGGCTGCGATA 1323
QY 1873 ttgttaatgcacagagattgaaccttcggtcttaacagactatgctatgcaatat 1932
Db 1324 TTGTTAATGCCATCCAGATTCGAACCTTGTGGTCTCAATCAGCTATATGCTATGAGTAT 1383
QY 1933 ggtacagttctctgattgctgaactggggccctccgagacacagctcagagacctcaac 1992
Db 1384 GGCACAGTTCCTGTTGCTCATGCAACTGGGGGCTTTAGAGATACCGTGGAGAACTTCAAC 1443
QY 1993 ccttttgggtcaaaagagagaggttacaggtggcggttctccaccgctaaacgtggac 2052
Db 1444 CCTTCGGTGAGAAATGGAGAGCAGGTACAGGGTGGGCAATTCGACCCCTAACACAGAA 1503
QY 2053 aagatg-ttgtgggcatgccaaccgcatgctgacattcagggagcacaagcgcctctg 2111
Db 1504 AACATGTTGTGACATTCGCACTGCAATATCTACATACAGGGAACAAAGTCCCTCTG 1563
QY 2112 ggaagggctaatgaagcggagcagatgacgaaagaccatacgtgggaccatgcccc 2165
Db 1564 GGAAGGGCTAATGAAGCGGAGCATGCAAAAGACTTTCACGTGGGACCATGCCGC 1617

```

RESULT 15

```

US-09-345-214-6/c
; Sequence 6, Application US/09345214
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Lightner, Jonathan E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; FILE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20

```

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: Zea mays
US-09-345-214-6

Query Match      43.9%; Score 1215.8; DB 17; Length 1528;
Best Local Similarity 87.7%; Pred. No. 9.4e-103;
Matches 1339; Conservative 0; Mismatches 187; Indels 1; Gaps 1;

QY 678 tcaagctaaagtacacgtagcatcgtttgtgactggtgaagctgctcccttatgcaaa 737
Db 1528 TCGAGGCTAAAGTAACACAAAGCATTTGCTTTGTAACCGCGGAAGCTTCTCCTTATGCAAA 1469
QY 738 gtcaggggggggttggagatggtttgtggttcgtttaccaaatgtctctgtcgtcgtggtca 797
Db 1468 GTCTGGGGGTCTAGGAGATGTTTGTGGTTCATTTGCCAGTTGCTTTCGTCTCGTGGTCA 1409
QY 798 ccgagtgatggttgaatcccaagatacttaaatgggtcctctctgataaaaaactatgcaaa 857
Db 1408 CCGTGTGATGGGTGTAATGCCAGATATTTAAATGTACTCTCGATAAGAAATTAATGCAAA 1349
QY 858 ggcattatcacactgcgaagcacatcaaatgattccatgcttttgggggacacatgaagtgc 917
Db 1348 TGCATTTTACACAGAAAAACACATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTAC 1289
QY 918 ctttttcatgagtatagagacaacgtcagattgggtgttgttcgatcatccgtcatatca 977
Db 1288 CTTCTTCCATGATAGTATAGAGATTTCATTGACTGGGTGTTTGTGTGATCATCCCTCATATCA 1229
QY 978 cagaccagaaagtttatatggagataaattttggtgcttttgggtgataatcagttcagata 1037
Db 1228 CAGACCTGGAATTTATATATGGAGATAAAGTTTGGTGTCTTTTGGTCAATAACAGTTTCA 1169
QY 1038 cacactcctttgctatgctgcatgcgagggcccaactaaaccttgaattgggaggatata 1097
Db 1168 CACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTTGATCCTTGAATTTGGAGGATATAT 1109
QY 1098 tiatgagacaagattgctattgttgaacgattggcattggcagctcgtcgtcagctcct 1157
Db 1108 TTATGACAGAAATTTGATGTTTGTGTCATGATTGGCATGCCAGTCTAGTGCCAGTCT 1049
QY 1158 tcttgctgcaaaatatagaccatacgggttttacagagattcccgagcaccccttcttat 1217
Db 1048 TCTTGTGTCAAATATAGACCATATGTTGTTTATAAAGACTCCCGCAGCATTTCTGTAAT 989
QY 1218 acataatttagcacatcaggggtgtggagcctgcaagtacatatccttgatctggtggtgc 1277
Db 988 ACATAATTTAGCACATCAGGGTGTAGAGCTTCAAGCACATATCCTTGACCTTGGGTGCC 929
QY 1278 tctggaatggtatggagctttagaatgggtatttccagaatgggcaagagagcagctccct 1337
Db 928 ACCTGAATGATGGAGCTCTGGAGTGGGTATTTCCTTGAATGGCGGAGGAGGATGCCCT 869
QY 1338 tgacaaggggtgagcgagttlaactttttgaaaggagcaggttgtgacagcagatcggttgt 1397
Db 868 TGACAAGGGTGGAGCATTAATTTTGAAGGTGCAGTTGTGCACAGCAGATCGAATCGT 809
QY 1398 gaccgtcagtcaggggttatcatcagggaggttcacaactgctgaaggtggacagggcctcaa 1457
Db 808 GACTGTCAAGTAAAGGTTATTTCGTGGGAGGTCAACAACCTGCTGAAGGTGGACAGGCGCTCAA 749
QY 1458 tgaactcttaagctcccgaaaaagtattgaatgaattgaaatgaaatgaaatgaaatgaaat 1517
Db 748 TGAGCTCTTAAAGCTCCAGAAAGAGTGTATTAAACGGAATTTGTAATGGAATTAATGCAAT 689
QY 1518 tgattggaaccccccacacagacagaaagttctccctcatcattattctgtcgtacacctctc 1577
Db 688 TGATTGGAACCCCTGCCACAGACAAATGATCCCCCTGCTATTATTTCTGTGATGACCTCTC 629
QY 1578 tggaaagggccaaatgtaaagctgaattgcagaagggaggttgggtttacctgttaaggagga 1637

```

```

|||||
Db 628 TGGAAAGGCCAATGTAAGGTGCTATTCAGAGAGAGCTGGGTTTACCTATAAGGCCTGA 569
Qy 1638 tgttccctctgattggtcctttattggaagactggattaccagaaagcatttgatctcattaa 1697
Db 568 TGTTCCTCTGATTGGCTTTATTGGAAGGTTGGATTATCAGAAAGGCATTGATCTCATTTCA 509
Qy 1698 aatggccattccagagatcattgagggagcgctgcaatttgatcgttgatctgggga 1757
Db 508 ACTTATCATACCAAGATCTCATCGGGAAGATGTTCAATTTGTCATGCTGGATCTGGTGA 449
Qy 1758 tceaattttgaagctggatgagatcaccgagtcgagttacaagataaaattccgtgg 1817
Db 448 CCCAGAGCTTGAAGATTGGATGAGATCTACAGATCGATCTTCAAGGATAAATTTCTGGTGG 389
Qy 1818 atgggttggatttagtgttccagtttcccccacagaataactgcaggttgcgatatattgtt 1877
Db 388 ATGGGTTGGATTAGTCTTCCAGTTTCCCAACCGAATAACTGCCGGCTGCCGATATATTGTT 329
Qy 1878 aatgceatcgagatttgaaccttgcggtcttaatacagctatgctatgcaaatatggtac 1937
Db 328 AATGCCATCCAGATTCCGAACCTTTGTGCTCTCAATCAGCTATATGCTATGAGTATGGCAC 269
Qy 1938 agttccctgtagttcatggaactggggcctcccgagacacagtcgagaccttcaaccttt 1997
Db 268 AGTTCTCTGTTGTCCATGCAACTGGGGGCCCTTAGAGATACCGTGGAGAACTTCAACCCCTT 209
Qy 1998 tgggtcaaaaggagagaggtgtacaggggtggcggttctcaccgcctaaacctggacaagat 2057
Db 208 CGGTGAGAATGGAGAGCAGGGGTACAGGGTGGGCATTGGCACCCCTAACCCACAGAAACAT 149
Qy 2058 gttgtgggcatttcgaaaccgagatgtcgcacattcagggagcacaagccgtccctgggaggg 2117
Db 148 GTTGTGGACATTGGCAACTGCAATATCTACATACAGGGAAACACAAGTCTCTCTGGGAAGG 89
Qy 2118 gctcatgaagcgggcatgacgaagaccatacgtgggaccatgcc-ccgagcagtacga 2176
Db 88 GCTAATGAACGAGGCATGTCAAAAGACTTCACGTGGGACCATTGCCGCTGAACAATACGA 29
Qy 2177 gcagatcttcgagtgggccttcgtgga 2203
Db 28 ACAAACTTCCAGTGGGGCCTTCATCGA 2

```

Search completed: March 28, 2002, 16:19:36
Job time: 9070 sec

Result No.	Score	Query Match	Length	DB	ID	Description
1	2771	100.0	2771	5	US-09-674-824-1	Sequence 1, Appli
2	2541.8	91.7	2662	8	US-09-508-377-11	Sequence 11, Appl
3	2239	80.8	2239	7	US-09-952-677-11	Sequence 1, Appli
4	1393	50.3	2383	7	US-09-931-297-1	Sequence 1, Appli
5	567	20.5	10337	8	US-09-508-377-13	Sequence 13, Appli
6	540	19.5	660	8	US-09-873-402A-83353	Sequence 83353, A
7	399	14.4	589	8	US-09-873-402A-67388	Sequence 67388, A
8	355.2	12.8	561	7	US-09-865-439A-118661	Sequence 118661, A
9	307.8	11.1	407	6	US-09-985-678-88686	Sequence 88686, A
10	307.6	11.1	411	6	US-09-985-678-83476	Sequence 83476, A
11	299	10.8	598	9	US-10-021-323-969	Sequence 969, App
12	299	10.8	5072	8	US-09-508-377-15	Sequence 15, Appli
13	293.6	10.6	412	7	US-09-865-439A-21857	Sequence 21857, A
14	293.6	10.6	635	7	US-09-596-564A-17004	Sequence 17004, A
15	286.6	10.3	577	7	US-09-849-529A-17365	Sequence 17365, A
16	279	10.1	335	6	US-09-985-678-9898	Sequence 9898, Ap
17	272.4	9.8	2418	9	US-10-044-543-25	Sequence 25, Appli
18	269.4	9.7	372	7	US-09-865-439A-29137	Sequence 29137, A
19	255.8	9.2	316	6	US-09-985-678-39564	Sequence 39564, A
20	255	9.2	2348	9	US-10-044-543-5	Sequence 5, Appli
21	251.6	9.1	951	5	US-09-819-091A-6343	Sequence 6343, Ap
22	249.2	9.0	286	6	US-09-985-678-41547	Sequence 41547, A
23	237.4	8.6	289	6	US-09-985-678-90311	Sequence 90311, A

QY 241 cctctccgctctccctctcccgccgcccacacccatgcgcgacggcgctggc 300
DB 241 cctctccgctctccctctcccgccgcccacacccatgcgcgacggcgctggc 300
QY 301 gcggggtgctccgccccagcgtccgctcgccgccccgccccgccccgccccg 360
DB 301 gcggggtgctccgccccagcgtccgctcgccgccccgccccgccccgccccg 360
QY 361 tccgctcgctccgccccgccccgccccgccccgccccgccccgccccgccccg 420
DB 361 tccgctcgctccgccccgccccgccccgccccgccccgccccgccccgccccg 420
QY 421 ctaagcaggagggccccgccccgccccgccccgccccgccccgccccgccccg 480
DB 421 ctaagcaggagggccccgccccgccccgccccgccccgccccgccccgccccg 480
QY 481 gtgcaggcttctccgccccgccccgccccgccccgccccgccccgccccgccccg 540
DB 481 gtgcaggcttctccgccccgccccgccccgccccgccccgccccgccccgccccg 540
QY 541 ccgccccgccccgccccgccccgccccgccccgccccgccccgccccgccccg 600
DB 541 ccgccccgccccgccccgccccgccccgccccgccccgccccgccccgccccg 600
QY 601 gctgagattccatcgacagataattgtggctgcaagtgcagcagattctgcagatcg 660
DB 601 gctgagattccatcgacagataattgtggctgcaagtgcagcagattctgcagatcg 660
QY 661 gatcggaatgagcaacctcaagctaaagttaacgctgacgctgctgtgtgactggta 720
DB 661 gatcggaatgagcaacctcaagctaaagttaacgctgacgctgctgtgtgactggta 720
QY 721 gctgctctctcaaaagtgcgggggtgggagatgttctggtctgcttaccaattgct 780
DB 721 gctgctctctcaaaagtgcgggggtgggagatgttctggtctgcttaccaattgct 780
QY 781 ctgtgctgctgtgtaacggagtgatggtgtaatagcgaagatacttaaatgggtctct 840
DB 781 ctgtgctgctgtgtaacggagtgatggtgtaatagcgaagatacttaaatgggtctct 840
QY 841 gataaaactatgaaaggcattatacactgcgaagcacataagattccatgcttggg 900
DB 841 gataaaactatgaaaggcattatacactgcgaagcacataagattccatgcttggg 900
QY 901 ggatacatgaagtgcaccttttctatgagtatagagacaacgtcgattgggtttgtc 960
DB 901 ggatacatgaagtgcaccttttctatgagtatagagacaacgtcgattgggtttgtc 960
QY 961 gatactcgctatatacagaccaggaaatttatatgagataatttgggtgctttgg 1020
DB 961 gatactcgctatatacagaccaggaaatttatatgagataatttgggtgctttgg 1020
QY 1021 gataatcagttcagatacacactcttctgctatgctgcatgcgagccccactaatcct 1080
DB 1021 gataatcagttcagatacacactcttctgctatgctgcatgcgagccccactaatcct 1080
QY 1081 gaattggaggataatttatggacagaattgcatgttgggttgaaagattggcatgcc 1140
DB 1081 gaattggaggataatttatggacagaattgcatgttgggttgaaagattggcatgcc 1140
QY 1141 agcctgtgcaagctcttctgctgcaaaatagaccatacgggtttacagagattcc 1200
DB 1141 agcctgtgcaagctcttctgctgcaaaatagaccatacgggtttacagagattcc 1200
QY 1201 cgcagcaccttctgtatataataatttagacatcagggtgtggagcctgcaagtacat 1260
DB 1201 cgcagcaccttctgtatataataatttagacatcagggtgtggagcctgcaagtacat 1260
QY 1261 cctgactggggttgcctcctgaaaggatggagcttttagaattgggtattccagaaagg 1320
DB 1261 cctgactggggttgcctcctgaaaggatggagcttttagaattgggtattccagaaagg 1320
QY 1321 gcaaggagcgtcccttgacaaagggtgaggcgagttacttttgaaaggagcagttgtg 1380

DB 1321 gcaaggagcgtcccttgacaaagggtgaggcagttacttttgaaaggagcagttgtg 1380
QY 1381 acagcagatcgagattgtaccgctcagtggtttattcattggagggtcacaaactcgtgaa 1440
DB 1381 acagcagatcgagattgtaccgctcagtggtttattcattggagggtcacaaactcgtgaa 1440
QY 1441 ggtggacagggcctcaatgagctcttaagctcccccgaagtgattgaatggaaattgta 1500
DB 1441 ggtggacagggcctcaatgagctcttaagctcccccgaagtgattgaatggaaattgta 1500
QY 1501 aatgggaattgacatttaattggaacccccacacagacaagtgtctccctcatcattat 1560
DB 1501 aatgggaattgacatttaattggaacccccacacagacaagtgtctccctcatcattat 1560
QY 1561 tctgctcagatgaactctctggaagggccaaatgttaaagctgaattgcagaaaggagttgggt 1620
DB 1561 tctgctcagatgaactctctggaagggccaaatgttaaagctgaattgcagaaaggagttgggt 1620
QY 1621 ttacctgaaggagagatttctctctgattggtctttatttgggaagactggaattaccagaaa 1680
DB 1621 ttacctgaaggagagatttctctctgattggtctttatttgggaagactggaattaccagaaa 1680
QY 1681 ggcatgactcattaaaaatggccattccagagctcattgaggagagacgtgcaatttgc 1740
DB 1681 ggcatgactcattaaaaatggccattccagagctcattgaggagagacgtgcaatttgc 1740
QY 1741 atgcttggaactctgggagccaaatcttctggaaggctggaattccacagttccacagaataactgca 1800
DB 1741 atgcttggaactctgggagccaaatcttctggaaggctggaattccacagttccacagttac 1800
QY 1801 aaggataaaatccogtggattgggtttagtctccagtttccacagataactgca 1860
DB 1801 aaggataaaatccogtggattgggtttagtctccagtttccacagataactgca 1860
QY 1861 ggttgcgatataattgtaatagcacatcgagatttgaaccttgcggtcttaactcagctat 1920
DB 1861 ggttgcgatataattgtaatagcacatcgagatttgaaccttgcggtcttaactcagctat 1920
QY 1921 gctatgcaaatggttacagttctctgttagttcaatggaaactgggggctcccgagacacagtc 1980
DB 1921 gctatgcaaatggttacagttctctgttagttcaatggaaactgggggctcccgagacacagtc 1980
QY 1981 gagacctcaaccttttctgcaaaaggagagaggttacaggggtggcggttctcaccg 2040
DB 1981 gagacctcaaccttttctgcaaaaggagagaggttacaggggtggcggttctcaccg 2040
QY 2041 ctaacccgtgacaaagattgttggggcattgcgaacccgcgattgcgaattcaggggagcac 2100
DB 2041 ctaacccgtgacaaagattgttggggcattgcgaacccgcgattgcgaattcaggggagcac 2100
QY 2101 aagccgtctctgggaggggctcattgaagcaggcatgacgaagaccatacgtgggacct 2160
DB 2101 aagccgtctctgggaggggctcattgaagcaggcatgacgaagaccatacgtgggacct 2160
QY 2161 gccccgagcagtagcagcagatcttcagtgggccttcgtggacaaacctacgtcattgt 2220
DB 2161 gccccgagcagtagcagcagatcttcagtgggccttcgtggacaaacctacgtcattgt 2220
QY 2221 agacgggagctggggaggtccaaatgcgagctcctctcagctctgaagacatcctctca 2280
DB 2221 agacgggagctggggaggtccaaatgcgagctcctctcagctctgaagacatcctctca 2280
QY 2281 tccctcgcggcccggaagatacccttgacattgcgttgcgtgctacagtagagtcg 2340
DB 2281 tccctcgcggcccggaagatacccttgacattgcgttgcgtgctacagtagagtcg 2340
QY 2341 caatgcgctgcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 2400
DB 2341 caatgcgctgcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 2400
QY 2401 gcggtgacagcttcgggtgcgacagcttaccagtttggggaataaggaaggagtgctg 2460

Db	2401	gcggtacagcttcgggtgcagcagttacagttttggggaataaggagggtgtgct	2460
Qy	2461	gcaggatggttaacagcaagcaccactcagatggcagctctctctgtccgtgttacagct	2520
Db	2461	gcaggatggttaacagcaagcaccactcagatggcagctctctctgtccgtgttacagct	2520
Qy	2521	gaatcagaaaaaccacgctgactcttttagccttagttagtgtaaagttttgttccttct	2580
Db	2521	gaatcagaaaaaccacgctgactcttttagccttagttagtgtaaagttttgttccttct	2580
Qy	2581	gtgtatgttcttcttccttagctgacaaattattgacctgttgagagaattttatctttg	2640
Db	2581	gtgtatgttcttcttccttagctgacaaattattgacctgttgagagaattttatctttg	2640
Qy	2641	ctctgttttttttaatacaagggggtttctccgatttcatataaaaaaataaaaa	2700
Db	2641	ctctgttttttttaatacaagggggtttctccgatttcatataaaaaaataaaaa	2700
Qy	2701	aaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa	2760
Db	2701	aaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa	2760
Qy	2761	aaaaaaaaaaaa 2771	
Db	2761	aaaaaaaaaaaa 2771	

RESULT 2

US-09-508-377-11

; Sequence 11, Application US/09508377

; GENERAL INFORMATION:

; APPLICANT: KALEEN, ZHONGYILI

; APPLICANT: MORELL, MATTHEW

; APPLICANT: RAHMAN, SADEOUR

; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS

; FILE REFERENCE: 054270/0126

; CURRENT APPLICATION NUMBER: US/09/508, 377

; CURRENT FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: AU PP 2509

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: PCT/AU98/00743

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: AU PP 9108

; PRIOR FILING DATE: 1997-09-12

; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 2662

; TYPE: DNA

; ORGANISM: Triticum tauschii

US-09-508-377-11

Query Match . 91.7%; Score 2541.8; DB 8; Length 2662;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2597; Conservative 0; Mismatches 32; Indels 6; Gaps 3;

Qy	25	ctcccactcttctctcccgcgcacaccgagtcggcaccggtcatcaccatcacctcg	84
Db	2	ctcccactcttctctcccgcgcacaccgagtcggcaccggtcatcaccatcacctcg	61
Qy	85	gcctcggccaccggcaaaccccccgatcgcgttttcgaggcagcgactaaaaacccggg	144
Db	62	gcctcggccaccggcaaaccccccgatcgcgttttcgaggcagcgactaaaaacccggg	121
Qy	145	gagcgcgccccgcggcagcagcagcaccgcactggagagagagaggttcgcccgaccgcg	204
Db	122	gagcgcgccccgcggcagcagcagcaccgcagtcggaggagagaggttcgcccgaccgcg	181
Qy	205	caccgagcggggggatccaccgctccgttgcgctcgcacactctccgcctctcccctgtcc	264
Db	182	caccgagcggggggatccaccgctccgttgcgctcgcacactctccgcctctcccctgtcc	241

Db 1322 ggtgagcagtttaacttttgaaaggagcagtcgtgacagcagatggaattgacgcgc 1381
Qy 1405 agtcagaggttatcatggtgaggtcacaaactgctgaaagtgtgacagggccctcaatgaagctc 1464
Db 1382 agtcaggggttatcatggtgaggtgacaaactgctgaaagtgtgacagggccctcaatgaagctc 1441
Qy 1465 ttaagctcccgaaaaaagctgattgaaatggaaatgttaaatggaattgacattaatgattgg 1524
Db 1442 ttaagctcccgaaaaaagctgattgaaatggaaatgttaaatggaattgacattaatgattgg 1501
Qy 1525 aacccacacacagagaagtgctccctcatcattattctgtcgatgacacctctctgtgaaag 1584
Db 1502 aacccacacacagagaagtgctccctcatcattattctgtcgatgacacctctctgtgaaag 1561
Qy 1585 gcaaaatgtaaaagctgaattgcagaagagtggtgttacctgttaaggaggaggtgttcct 1644
Db 1562 gcaaaatgtaaaagctgaattgcagaagagtggtgttacctgttaaggaggaggtgttcct 1621
Qy 1645 ctgattggctttattggaagactggattaccagaagaagcattggtatctcattaaaatggcc 1704
Db 1622 ctgattggctttattggaagactggattaccagaagaagcattggtatctcattaaaatggcc 1681
Qy 1705 attccagagctcatgagggagagcgtcaatttgcattgtctgattggtggtggtcccaatt 1764
Db 1682 attccagagctcatgagggagagcgtcagcttgcattgtctgattggtggtggtcccaatt 1741
Qy 1765 ttgaaaggctggatgagatctaccagagtcagagttacaagataaattccgtggatgggtt 1824
Db 1742 ttgaaaggctggatgagatctaccagagtcagagttacaagataaattccgtggatgggtt 1801
Qy 1825 ggatttgggttccagtttccacacagataaactcaggttgcgatatattgttaaatgcca 1884
Db 1802 ggatttgggttccagtttccacacagataaactcaggttgcgatatattgttaaatgcca 1861
Qy 1885 tcgagatttgaaccttgcgttcttaacagctatatctatgcaaatggtcacagttcct 1944
Db 1862 tccagggttgaaaccttgcgttcttaacagctatatctatgcaaatggtcacagttcct 1921
Qy 1945 gtatttcatggaactgggggctccgagacacagtcgagaccttcaaccttttgggtgca 2004
Db 1922 gtatttcatggaactgggggctccgagacacagtcgagaccttcaaccttttgggtgca 1981
Qy 2005 aaaggagagaggttacaggttgggcttccacgctaaacctggaagaagatttgg 2064
Db 1982 aaaggagagaggttacaggttgggcttccacgctaaacctggaagaagatttgg 2041
Qy 2065 gcattggaaccgcatgctgacattcaggagacacagccgtctctggaggggctcatg 2124
Db 2042 gcattggaaccgcatgctgacattcaggagacacagccgtctctggaggggctcatg 2101
Qy 2125 aagcagaggtcatgcaaaagaccatactgtgggaccatgcc-ccgagcagttacgacagatc 2183
Db 2102 aagcagaggtcatgcaaaagaccatactgtgggaccatgcgcgcgagcagttacgacagatc 2161
Qy 2184 ttgaggtgggcttctgtgagacaaacctacgtcatgtagacgggagctgggaggtccaa 2243
Db 2162 ttgaaatgggcttctgtgagacaaacctacgtcatgtagacgggagctgggaggtcgaa 2221
Qy 2244 gtgcgagctctcttcagctctgaagacatctcttcatcttccgctccgcccgggaagata 2303
Db 2222 gcgcgggctctcttgagctctgaagacatgttctcatcttccgctccgcccgggaagata 2281
Qy 2304 cccctgtacatttgcgttctgtctacagtagagtcgcaatgcgcctctgttcttgggt 2363
Db 2282 cccctgtacatttgcgttctgtctacagtagagtcgcaatgcgcctgtctgc-ttggtc 2340
Qy 2364 gcgcgggttcgagacatatagcgcgtgtgtctgtcgcggcgtgacagcttcgggtggagc 2423
Db 2341 gcgcgggttcgagagtagatagcgcgtgtgtctgtcgcggcgtgacagcttcgggtggagc 2400
Qy 2424 acagttacagtttggggaaataaggaaaggaggtgtgtctgcagatggtttaacagcaagca 2483

Db 2401 acagttacagtttgggggaataaaggaggaggtgtgtgcagaggtggttaacagcaagca 2460
Qy 2484 ccactcagatggcagcctctctgtccgtgttacagctgaaatcagaacccaactggtgac 2543
Db 2461 ccactcagatggcagcctctctgtccgtgttacagctgaaatcagaacccaactggtgac 2520
Qy 2544 tccttagccttagtgattgtgaagtttggccttctgtgtatgtatgttgcctttagc 2603
Db 2521 tccttagccttagtgattgtgaagtttggccttctgtgtatgtatgttgcctttagc 2580
Qy 2604 tgacaaatattgacctgttggagaat----tttattcttgcgtgctgtttttt 2654
Db 2581 tgacaaatattgacctgttggagaattttatttattcttgcgtgctgtttttt 2635

RESULT 3
US-09-952-677-1
: Sequence 1, Application US/09952677
: GENERAL INFORMATION:
: APPLICANT: Block, Martina
: Lorz, Horst
: Lutticke, Stephanie
: Walter, Lennart
: Froberg, Claus
: Kosmann, Jens
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
: FROM WHEAT WHICH ARE INVOLVED IN STARCH
: SYNTHESIS
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/952.677
: FILING DATE: 14-Sep-2001
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/196,390
: FILING DATE: 19-Nov-1998
: APPLICATION NUMBER: DE 196 21 588.9
: FILING DATE: 29-MAY-1996
: APPLICATION NUMBER: DE 196 36 917.7
: FILING DATE: 11-SEP-1996
: APPLICATION NUMBER: PCT/EP97/02793
: FILING DATE: 28-MAY-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Haley, Jr., James F.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: AGREVO-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 596-9000
: TELEFAX: (212) 596-9090
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2239 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Triticum aestivum L.
: STRAIN: cv. Florida
: HAPLOTYPE: ca. 21 d Caryopses


```

; IMMEDIATE SOURCE:
; LIBRARY: CDNA library in pBluescript sk (-)
; CLONE: TaSSS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1017
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-952-677-1

Query Match      80.8%; Score 2239; DB 7; Length 2239;
Best Local Similarity 100.0%; Pred. No. 2e-311;
Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 533 cgacgcagccgcctcccgagccgcgtgagggaactcgcgcgcgacctcctactcg 592
Db      |||
QY 533 aagggtatgtgaggttcccatcgacagcataattgtgctgcaagtggacagattctg 652
Db      |||
QY 61  AAGGGATTGTGAGGATTCCATCGACAGCATAATTGTGGCTGCAAGTGGACAGGATTCG 120
Db      |||
QY 653 agatcatggatgcaatgagcaacctcaagctaaagtacacgtagcatcggtctgtga 712
Db      |||
QY 121 AGATCATGGATGGAATGACAACTCAAGCTAAAGTTACACATAGCATCGTGTGTGA 180
Db      |||
QY 713 ctggtgaagctgctctctatgcaaaagtcaggggggtggagagatgttgggttcgttac 772
Db      |||
QY 181 CTGGTGAAGCTGCTCCTTATGCAAGTCAGGGGGGTGGGAGATGTTTGGTTCGTTAC 240
Db      |||
QY 773 caattgctctgtgctggtgtccaccgagtgatggtgttaatgccaaagatacttaaatg 832
Db      |||
QY 241 CAATTGCTGTGCTGCTGTGTGTCACCGAGTGATGTTGTTAATGCCAAGATACTTAAATG 300
Db      |||
QY 833 ggtcctctgataaaactatgcaaggcattatacactcgaagcacatgaagattccat 892
Db      |||
QY 301 GGTCTCTGATANAACACTATGCAAGGCATATACACTCGGAAGCACATTAAGATTCCAT 360
Db      |||
QY 893 gctttgggggatacacatgaagtaccttttctatgagtagatagagacaacgtcgtatgg 952
Db      |||
QY 361 GCTTTGGGGGATCACATGAAGTGACCTTTTTCATGAGTATAGAGACAACGTCGATTGG 420
Db      |||
QY 953 tgtttgcgatacatccgtcatatcacagaccaggagtttatatgagagataatttggcg 1012
Db      |||
QY 421 TGTTTGTCCATATCCGTCATATCACAGACCAGGAAGTTTATATGGAGATAATTGTGTG 480
Db      |||
QY 1013 cttttggtgataatcagttcagatacacactcctttgtctatgctcgtcagtcgagggccac 1072
Db      |||
QY 481 CTTTGTGTGATAATCAGTTTCAGATACACACTCCTTTGCTATGCTGCATCGGAGGCCAC 540
Db      |||
QY 1073 taatcctgaattggaggatattatgacagaaattgcattgttgttgaacgatt 1132
Db      |||
QY 541 TAATCCTTGAATGGGAGGATATTTATGAGACAAATTCATGTTGTGTGTAACGATT 600
Db      |||
QY 1133 ggcagtcgagccttggcagtccttctgtcgcaaatatagaccatacagtggtttaca 1192
Db      |||
QY 1193 gagattcccgacgacccttgttatatacaatttagcacatacaggtgtgagccctgcaa 1252
Db      |||
QY 661 GAGATTCCCGCAGCACCCTTGTATACATAATTTAGCACATCAGGGTGTGGAGCCTGCAA 720
Db      |||
QY 1253 gTacatactctgactgaggttgctcctgtaattgagcatttagagctttaaaatgggtatttc 1312
Db      |||
QY 721 GTACATATCCTGATCTGGGATGGCTCTGTAATGATGAGAGCTTTAGAATGGGTATTTC 780
Db      |||
QY 1313 cagaatggcagaggagccttgcacaaaggtgagggcaggttaactttttgaaagag 1372
Db      |||
QY 781 CAGAAATGGCAAGGAGGATGCCCTTGACAAGGGTGGAGCAGTTAACTTTTGAAGGAG 840
Db      |||
QY 1373 cagttgtacagcagatcggattgtgaccgtcaagtcagtcaggttattcatgaggagtcacaa 1432
Db      |||
QY 841 CAGTTGTGACAGCAGATCGGATTGTGACCCGTGAGTCAGGTTTATTCATGGGAGGTCACAA 900
Db      |||
```

```
QY 2513 ttacagctgaatcagaacccaactgtagctcttagcttagtgattgtaagttgt 2572
DB 1981 ttacagctgaatcagaacccaactgtagctcttagcttagtgattgtaagttgt 2040
QY 2573 tgcctctgtgtatgtgtcttcttagctgacaaaattttagcctgttgagaaattt 2632
DB 2041 TGCTCTCTGTGTATGTCTGTCTTACCTGACAAATATTGACCTCTTGGAGAATT 2100
QY 2633 tatcttctgtctgttttttttaatacaaaagagggttctctcgatttcatataaaa 2692
DB 2101 TATCTTTGCTGCTGTTTTTTTTTAAATCAAAAGAGGGGTTCTCTCGATTTTCA 2160
QY 2693 aaaaaaatacaaaagagggttctctcgatttcatataaaa 2752
DB 2161 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2220
QY 2753 aaaaaaatacaaaagagggttctctcgatttcatataaaa 2771
DB 2221 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2239

RESULT 4
US-09-931-297-1
; Sequence 1, Application US/09931297
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; starch syntheses from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/931,297
; FILING DATE: 16-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/192,909
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2383 base pairs
; TYPE: nucleotide
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; TISSUE TYPE: endosperm
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1950
```

```
; OTHER INFORMATION: /function= "starch synthesis"
; /product= "soluble starch synthase"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-931-297-1
```

```
Query Match 50.3%; Score 1393; DB 7; Length 2383;
Best Local Similarity 83.5%; Pred. No. 2.le-190;
Matches 1619; Conservative 0; Mismatches 310; Indels 10; Gaps 3;

QY 302 ccgggtgctcgcccccaagctccgcccgtcgccgagatccgcgagcgccgcccgggcgct 361
DB 39 CCTCGGCGGTGGGCGCGCGTGCCTTCTTCCCTCGCGGGGCCCTTGGCGCGCGCGCTCG 98
QY 362 ccgctgctgctcgccgcccgtccgcccgtccgcccgtccgcccgtccgcccgtccgccc 421
DB 99 GCGACGCGGCGCGCGCGCGCGCTCCAGCGGTGCTGCGCGCGCGGTGCTGCGCGGAGC 158
QY 422 tcagcaggaggggcccccgccgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcgccc 481
DB 159 TGACGAGGAGGGCGCGCGCGCGCGCGATGCCACCGCGCTGTGCGCGCGCGCGCTCG 218
QY 482 tgccagggttctcgccgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcg 541
DB 219 TGCGCGGCTTCTCTCG---CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275
QY 542 cgcctctgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcgccc 601
DB 276 CGCCCGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 329
QY 602 ctgaggattccatcgacagcataattgtgctgcaagtgcgagcagagattctcgagatcag 661
DB 330 CTGAAGGTTCCATCGATAACACAGTAGTTGTGSCAAGTAGCAAGATTCTGAGATTGTGG 389
QY 662 atcggaatgacaaacctcaagctaaagtacacgtacacgtctgttctgactggtgaag 721
DB 390 TTGGAAGGAGCAAGCTCGAGCTAAAGTAAACACAAACATTTCTTTGTAACATGGCGAAG 449
QY 722 ctgctctctatgcaaaagtcaggggggttgaggagatgttctggttctggttaccgaattgctc 781
DB 450 CTCTCTCTTATGCAAGTCTGGGGGTCTAGGAGATGTTGTGTGTTCTATTTGCCAGTTGCTC 509
QY 782 ttgctctcggttcacccgagtgatgttgaatgcaagatacttaattgggtctctctc 841
DB 510 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 569
QY 842 ataaaaactgcaaaaggcattatatacactgcgagcacattaaagattccatgctttgggg 901
DB 570 ATAGAATTTATGCAATGCAATTTACACAGAAAAACACATTCGGATTCCATGCTTTGGCG 629
QY 902 gatcacatgaagtgaaccttttttcagtagtatagagacaacgctcgattgggtgtttgtcg 961
DB 630 GTGAACATGAAGTTACCTTCTTCATGAGTATAGAGATTTCAGTTGACTGGGTGTTGTTG 689
QY 962 atacatcgctcatatcacagaccaggaagtttatatgagaataatttttgctgttttggtg 1021
DB 690 ATCATCCTCATATPCACAGACCTGGAAATTTATATGGAGATATTTTAAATGTTGCTTTGGTG 749
QY 1022 ataactggttcagatacacactcctttgtctatgctgcatcgagggcccccaactaatcttg 1081
DB 750 ATAATCAGTTTCAGATACACACTCTTTTGTCTATGCTGCTATGAGGCTCTCTTTGGTCTTG 809
QY 1082 aattgggagatatatttatggcagaaattgcagttgttgtgaaacgattggcagatccca 1141
DB 810 AATTGGGAGGATATATTTATGGACAGAAATTCATGTTTGTGTAATGATTTGGATGCCA 869
QY 1142 gcttctgcccagctctcttctgctgcaaaaatagaccatacgggtgtttacagagattccc 1201
DB 870 GTCAGTGGCCAGTCTCTTCTGCTGCAAAATATAGACCATATGGTGTGTTTATAAGACTCCC 929
QY 1202 gcagcacccttgttatataataattagcacatcagggtgtgtggagcctgcagatacatc 1261
DB 930 GCAGCATCTCTGTAAATACATAAATTTAGCACATCATCAGGCTGTAGAGCTTGCAGCACATATC 989
```

Qy	1262	ctgactcgggattgcctcctgaatggtatggagcttttagaatgggtatttcagaatggg	1301
Db	990	 CTGACCTTGGGTTGCCACCTGAATGCTATGGAGCTCTGGAGTGGGTATTCCTCTGAATGGG	1049
Qy	1322	caaggaggcatgccctctgacaagggtgaagcagtttaacttttgaaggagcagtttgtga	1381
Db	1050	 CGAGGAGCATGCCCTTGCACAGGGTGAAGCGTAGTAAATTTTGTAAAGAGTGCAGTTGTGA	1109
Qy	1382	cagcagatcggattgtgacgcctcagtcaggttcagggttatctatgggaggtcacaaactcgtgaag	1441
Db	1110	CAGCAGATCGAATCGTGACTGTGCTAGTAAGGGTTATTTCATGGGAGTCAACAATCCTCGAG	1169
Qy	1442	gtgacagggccctcaatgagctcttaagctcccgaaaaagtgtattgaatggaattgttaa	1501
Db	1170	GTGGACAGGGCCCTCAATGAGCTCTTAAGCTCCAGAAAAGAGTGTATTAAACGGAAATTGTAA	1229
Qy	1502	atggaattgacattaatgatgtgaaaccccaacacagacaaagtgtctccctcatcatatt	1561
Db	1230	ATGGAATTGACATTAATGATGTGAACCCCTGCCACAGACAAATGTATCCCTGTCAATTATT	1289
Qy	1562	ctgcgatgacctctctgaaaggccaaatgtaaagctgaattgcagaaggagtgtgggtt	1621
Db	1290	CTGTTGATGACCTCTCTGAAAGGCCAAATGTAAAGGTGTCATTGCAGAAAGAGCTGGGTT	1349
Qy	1622	tacctgtaaaggaggagtgtctctctgattgggtctttattggaagactggattaccagaag	1681
Db	1350	TACCTATAAGCCCTGATGTTCTCTGATTGGCTTTATTGGAAGATTGSGATTATCAGAAAG	1409
Qy	1682	gcattgatctcattaaatggccattccagagctcatgagggagagcgtgcgaatttgtca	1741
Db	1410	GCATTGATCTCATCAACTATTATACCAAGATCTCATCGGGGAAGATGTTCATATTTGTCA	1469
Qy	1742	tgttggtatctggggatcccaatttttgaaggctggatgagactcaccgagtcagattaca	1801
Db	1470	TGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCA	1529
Qy	1802	aggataaattccgctggatgggttgatttagtttccagtttccacagaaataaactgcag	1861
Db	1530	AGGATAAATTCTGSGATGGGTGGATTAGTTGTTCAGTTTCCAGTTTCCACCGGAATAACTGCGG	1589
Qy	1862	gttgcatatatgtttaatgccatcgagatttgaaccttcggttcttaatcagctatatg	1921
Db	1590	GCTGCGATATATTGTTAATGCCATCCAGATTGCAACCTTGTGGTCTCAATCAGCTATATG	1649
Qy	1922	ctatgcaatatggttacagttctctgatttcattgaactggggcctccgagacacacagtcg	1981
Db	1650	CTATGCAGTATGGCACAGTCTCTGTGTTCATGTCAACTTGGGGGCTTTAGAGTACCGTGG	1709
Qy	1982	agacctcaaccttttgtgcaaaaggagaggggttacagggtggcgcttcttcaccgc	2041
Db	1710	AGAACTTCAACCTTTTCGTGAGAAATGGAGACGAGGGTACAGGGTGGGCATTTCGCACCCC	1769
Qy	2042	taacctggacaagatgtgtggccattgcgaaccgcagattgcacattcaaggagacaca	2101
Db	1770	TAACCACAAANACATGTTGGACATTCCGAATTCGAATATCTACATACAGGGAAACACA	1829
Qy	2102	agccgtctctggagggggtcatgaagogaggcatgacgaagaagccatacagttgggaccatg	2161
Db	1830	AGTCCTCTCTGGGAAGGGCTTAATGAAGCGAGGCATGTCAAAAGACTTCACGTGGGACCATG	1889
Qy	2162	cc-ccgagcagtacgacagacatcttcaggtgggcttcctcgaggacacccatcagtcagt	2220
Db	1890	CCGCTGAACAATACCAACAANATCTTCAGTGGGGCTTTCATCGATCGACCCCTATGTCTATGT	1949
Qy	2221	agacgggggaactggggaggt 2239	
Db	1950	AAAAAGGACCAAGTGGT 1968	

RESULT 5

RESULT 3
US-09-508-377-13

03 03 208-377-13
; Sequence 13, Application US/09508377


```

: PRIOR APPLICATION NUMBER: US 60/161,619
: PRIOR FILING DATE: 1999-10-26
: NUMBER OF SEQ ID NOS: 17472
: SEQ ID NO 17004
: LENGTH: 635
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: LIB3279-2
US-09-696-664A-17004

```

Query Match	10.6%	Score 293.6	DB 7	Length 635
Best Local Similarity	85.0%	Pred. No. 3.3e-33		
Matches 340	Conservative 0	Mismatches 59	Indels 1	Gaps 1
QY 1825	ggattaggttcagattccacagaaactcaggttcgatatattttaatgcc 1884			
Db 23	ggattaggttcagattccacacgaataactcggcgcgatatattttaatgcc 82			
QY 1885	tcgagattgaaccttgcggtcttaacagctatatgctatgcaatatggcacagttcct 1944			
Db 83	tcgagattgaaccttggtctcaatcagctatatgctatgcatgagtcggcacagttcct 142			
QY 1945	gtagttcatgaaactggggcctccgagacacagtcgagaaactccaaccttttggtgca 2004			
Db 143	gtgtccatgcaactggggccttagagataccgtggagaaactccaacctttcgggtgag 202			
QY 2005	aaaggagagagggtacagaggtggcgctctcacccgtcaacccgtggacaagatgtgtgg 2064			
Db 203	aatggagagcaggtacagaggtggcgattcgacccctaaccacagaaacatgtgtgg 262			
QY 2065	gcattggaaaccgcgattgcgacattcaggagacaaagccgtcctcctggaggggctcatg 2124			
Db 263	acattgcgaactgcaatctacatacaggaacacacagtcctcctcctggaggggctaatg 322			
QY 2125	aagcagagcgtatgcgaagaaccatacgttggaccatgcc-ccgagcagtcagacagcagtc 2183			
Db 323	aagcagagcagtcgaaagaagcttcacgtggaccatgccgtgaaacaatacgaacaacac 382			
QY 2184	ttcgagtgggccttcgtggaccaaacctcagtcagttaga 2223			
Db 383	ttcagatgggccttcactcgatgcacctatgtcatgtaaa 422			

Query Match 10.3%; Score 286.6; DB 7; Length 577;
Best Local Similarity 69.7%; Pred. No. 3.3e-32;
Matches 402; Conservative 0; Mismatches 174; Indels 1; Gaps 1;
QY 1166 caaaatagaccatcagtggtttacagagattccgcagcaccttggtatatacaatt 1225

```

RESULT      13
US-09-865-439A-21857
; Sequence 21857, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 21857
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3353-011-Pl-Kl-C1
US-09-865-439A-21857

```

	Query Match	10.6%	Score 293.6;	DB 7;	Length 412;
	Best Local Similarity	85.2%;	Pred. No. 3.2e-33;		
	Matches 351;	Conservative	0;	Mismatches 59;	Indels 2; Gaps 2;
DY	1203	cagcacctgtttatacataatttagcacataagggtgtggagcctgcgaagtacatatcc	1262		
Ddb	1	cagcatctctgtaatacataatttagcacatcttctgtgatagcctgcgaagcacatctgc	60		
DY	1263	tgatc-tgggattgccctccctgaatcggtatgagcttgtagaagggtatttcagaaatggg	1321		
Ddb	61	tggcggttggtgtgccacctgtatgtatgtaaactctgagctgggtattccctgaaatggg	120		
DY	1322	-caaggaggcatgcccttgacaagaagggtgaggcagctaactctttgaaaaggagcagttgtga	1381		
Ddb	121	cgaggaggcatgcccttgacaagggtgaggcagctaactctttgaaaagggtcagtgtga	180		
DY	1382	cagcagatcgattgtgaccgtcagtcagggttatctcatgggaggttcacaactcgtcgaag	1441		
Ddb	181	cagcagatcgaaatcgtgactgtcagtaagggttatctctgggttcacaactcgtcgtg	240		
DY	1442	gtgacaggccctcaatgaactcttaagctcccgaaaaaagtgatgtgaatggaaattgttaa	1501		
Ddb	241	gggacaggccctcaatgaactcttaagcttcagaaaggagtgattaaacgggaattgttaa	300		
DY	1502	atggaattgacattaatgattggaacccccaccacagacaagtgctccctcatcatta-t	1560		
Ddb	301	atggaattgacatttatgatttggaacccctgtcacagactaatgtatactctgtcattatt	360		
DY	1561	tctgtcatgaactctcttgaaaggccaatgttaaagctgaaattgcagaagg	1612		
Ddb	361	tctgtgtgatgaccttttggaaaggccaatgttaaaggtgcctatgcagaagg	412		

RESULT 14
: Sequence 17004, Application US/09696664A-17004
: GENERAL INFORMATION:
: APPLICANT: Abad, Mark S.
: APPLICANT: Andersen, Scott E.
: APPLICANT: Dubois, Patrice
: APPLICANT: Mahadeo, Debbie A.
: APPLICANT: Masucci, James D.
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(51721)B
: CURRENT APPLICATION NUMBER: US/09/696, 664A
: CURRENT FILING DATE: 2000-10-25

```

Db      1  ccaaataccgtccatttgggtttataaagatgctcggagcgctccttgtaattcataacc 60
Qy  1226  tagcacatcagggtgtggagccctgcgaagtacatatcctgatctgggagctcctcctggaat 1285
Db      61  ttgcgatcaggagtggaacctgcagcaaccttttaagaaatttgggatgacctcaaacct 120
Qy  1286  ggtatggagcttttagaatgggtatttccagaaatgggcaaggagggcatgccccttgacaaag 1345
Db     121  ggtatggggcatggaaatgggtgtttccctacatgggcaaggacacatgctcttgacacag 180
Qy  1346  gtgaggcagtttaactttttgaaaggagcaggtgtgcacagcagatcggattgtgaccgtca 1405
Db     181  gagaggctgtcaattattctaaagggtgccattgtgacatctgcgaatactactaccgtta 240
Qy  1406  gtcagggttattcatgggaggtcacaaactctgaaggtgacagggcctcaatgagctct 1465
Db     241  gtaagggtctatgcatgggaatacaaaactgttgaagggtgatatggtttaaatcagctat 300
Qy  1466  taagtcccgaaaaagtgattgaatggaatgttaaatggaattgacattaatgattgga 1525
Db     301  taagcagtcggaggagtgcttctgactggaattacaaacggcattcatattgctgaaatggg 360
Qy  1526  accccacacagacaagtgctccctcatcattattctgtcgatgacctctctggaaaagg 1585
Db     361  atccatctcagatgagcatattgcttccattattctgtctctctctctctctctctctctct 420
Qy  1586  ccaaatgaaagctgaattgcagaaggagttgggtttaccctgtaaggaggaggtgttcttc 1645
Db     421  ttcattgcaagactgctctccacaggaaacttggttttccaattaaagcctgggtgtccct 480
Qy  1646  tgattggctttattggaagactggtattacc-agaaggcattgatctcatataaatggcc 1704
Db     481  tgatgggatttatcggaactggaactaccagaaggcattgaccttgatcactctgggca 540
Qy  1705  attccagagctcatgaggaggagcgtgcaatttgta 1741
Db     541  attccagaacttatggaagatgactacaaattgtaa 577

```

Search completed: March 28, 2002, 15:15:52
Job time: 7581 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 14:55:35 ; Search time 71.31 Seconds
(without alignments)
785.295 Million cell updates/sec

Title: US-09-674-824-2
Perfect score: 4044
Sequence: 1 MAATGVCAGCLAPSRLRAD.....SDGSLSVRTAEIRNLQVL 756

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	4044	100.0	756 21 AAY50818	Wheat soluble star
2	3612	89.3	671 19 AAW23937	Wheat soluble star
3	3348	82.8	647 20 AAY09004	Wheat starch solub
4	2749	68.0	649 19 AAW38218	Maize starch synth
5	2609.5	64.5	626 15 AAR51231	Soluble rice starch
6	2557.5	63.2	583 19 AAW56488	Zea mays soluble s
7	2548.5	63.0	583 19 AAW70894	Maize starch solub
8	2390.5	59.1	539 19 AAW58491	Zea mays starch sy
9	1718.5	42.5	459 17 AAR99540	Soluble starch syn
10	1205.5	29.8	698 19 AAW56487	ze mays soluble s
11	1203.5	29.8	804 19 AAW70892	Maize starch solub

12	1174.5	29.0	799 21 AAB37567	Wheat starch synth
13	1174.5	29.0	799 21 AAB37597	Wheat starch synth
14	1165	28.8	798 21 AAB37566	Wheat starch synth
15	1148.5	28.4	799 19 AAW23938	Wheat granule-boun
16	1120.5	27.7	597 21 AAB37568	Wheat starch synth
17	1068	26.4	669 19 AAW70893	Maize starch solub
18	1068	26.4	669 19 AAW56486	ze mays soluble s
19	877	21.7	609 19 AAW56485	Oryza sativa starch
20	876	21.7	606 13 AAR25476	Granule-bound star
21	869	21.5	502 13 AAR25474	Central fragment f
22	865	21.4	533 19 AAW56484	ze mays waxy gene
23	865	21.4	637 19 AAY16604	Protein encoded by
24	863	21.3	563 21 AAG04667	Arabidopsis thalia
25	863	21.3	610 21 AAG04666	Arabidopsis thalia
26	862.5	21.3	527 21 AAG04668	Arabidopsis thalia
27	769.5	19.0	209 21 AAG41355	Arabidopsis thalia
28	714.5	17.7	193 21 AAG41356	Arabidopsis thalia
29	678.5	16.8	477 13 AAR25462	giga. Escherichia
30	678.5	16.8	477 19 AAW53890	E. coli glycogen s
31	632.5	15.6	495 19 AAW70885	protein encoded by
32	617.5	15.3	277 19 AAY85849	S. pneumoniae deri
33	556	13.7	677 17 AAR99539	Soluble starch syn
34	556	13.7	1230 18 AAW17785	Potato tuber solub
35	556	13.7	1230 21 AAB49306	Potato starch synt
36	551	13.6	735 18 AAW27229	z. mays starch syn
37	511.5	12.6	534 21 AAB49307	Wheat starch synth
38	498.5	12.3	1059 21 AAB37570	Wheat starch synth
39	495.5	12.3	1828 21 AAB37569	Wheat starch synth
40	495.5	12.3	1628 21 AAB49304	Wheat starch synth
41	492	12.2	1674 20 AAY06199	Maize endosperm st
42	492	12.2	1674 21 AAG41354	Arabidopsis thalia
43	469.5	11.6	265 21 AAG41354	Arabidopsis thalia
44	448.5	11.1	1170 22 AAB70781	Maize starch synth
45	431.5	10.7	293 22 AAU00036	Wheat Starch synth

ALIGNMENTS

RESULT 1
AAY50818
ID AAY50818 standard; Protein; 756 AA.
XX
AC AAY50818;
XX
XX
DT 18-FEB-2000 (first entry)
XX
DE Wheat soluble starch synthase protein.
XX
KW Soluble; starch synthase; wheat; transgenic plant; starch production;
KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;
KW adhesive; textile; building material; soil stabilizer; wetting agent;
KW fertilizer; plant-protection; cosmetic; flocculant.
OS Triticum aestivum.
XX
XX DE19820607-AL.
XX
PD 11-NOV-1999.
XX
XX 08-MAY-1998; 98DE-1020607.
XX
PR 08-MAY-1998; 98DE-1020607.
XX
PA (AGRE) HOECHST-SCHERING AGREVO GMBH.
XX
PI Loerz H, Luetticke S, Block M;
XX
XX WPI; 2000-024508/03.
DR N-PSDB; AA224487.
XX
PT New enzyme with starch synthase activity, useful for producing starch
for foods and packaging materials -

XX Claim 1a: Page 19-21: 24pp: German.
XX This invention describes a novel protein (I) with the activity of wheat
CC starch synthase. Transgenic plants, specifically wheat, that contain (I)
CC are used for production of starch, used particularly in foods,
CC particularly baked and pastry goods and for making packaging materials or
CC disposable items. Starch may also be used as starting materials for
CC glucose or glucan components (e.g. for fermentation or further chemical
CC conversion); in paper and pulp production, as adhesives, in textiles,
CC in preparation of gypsum-based building materials, as soil stabilizer,
CC as wetting agent etc. in fertilizer and plant-protection compositions,
CC as binder (in pharmaceuticals, cosmetics, coal briquetting and casting
CC sand), as flocculant in soil or coal slurries, as rubber and leather
CC additives, and for production of synthetic polymers, e.g. polyurethane
CC films. Transgenic plants with increased/decreased production of (I)
CC produce starches with altered physical and/or chemical properties such as
CC amylose/amylopectin ratios, degree of branching, mean chain length,
CC phosphate content, gelatinization properties, gel- or film-forming
CC properties, or starch grain size or structure. This sequence represents
CC the soluble starch synthase isolated from wheat (Triticum aestivum L. cv.
XX Florida).
XX Sequence 756 AA;

Query Match 100.0%; Score 4044; DB 21; Length 756;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAATGVCAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSEGGPAARPAQ 60
Db 1 maatgvagcclapsvrladpatacaarasacvvrarlrargryvaelsregpaarpaqg 60
QY 61 QQLAPPLVPGFLAPPAPAPQSPAPTQPLPDAGVAGELAPDLLLEGIAEDSIDSIIVAAAS 120
Db 61 qqlapplvpgflappapapqspaptqplpdagvagelapdlllegiaedsidsiivaas 120
QY 121 EQDSEIMDANEQPAQKATRSIVFTGEAPYAKSGGLGVCGSLPTALAAARGHRVWVMP 180
Db 121 eqdseimdanegpqakvtrsvftgeapayaksgglgvcgsplalaaarghrvwnvmp 180
QY 181 RYLGSSDKNYAKALYAKHAKIKPCFGSGHEVTFHEYRDNDVWFVDHPSYHRPGSLYG 240
Db 181 rylngssdknyakalyakhakikpcfgsghevtffheyrdndvdfvhdpsyhrpgslyg 240
QY 241 DNFAGCDNQFRTLICYAACAPLILELGGYIYGONCMFVNDWHASLVVLLAAKYRP 300
Db 241 dnfgagdqnfrytllicyaaceaplilelgyiygqncmfvndwhaslvpvllaakyrp 300
QY 301 YGVYRDSRLTVLTHNAHQGVPEASTYPDGLPPEYNGALEWVFPWARHDLKGEAVN 360
Db 301 ygvyrdsrslvtlthnahogvpeastypdglppeyngalewvfpwarhaldkgeavn 360
QY 361 FLKGAUVVADRIVTVSQGYSWEVTTAEGGOGNELLSSRKSVLNGIVNGIDINDWNPSTD 420
Db 361 flkgavvtadrivtvsgqyswevttaeggggnellssrksvlingvngidindwnpttd 420
QY 421 KCLPHYSVDDLSGKAKCAELQELGLPVREDVPLIGFTRLDYQKGLDLKMAIPELM 480
Db 421 kclphysvddlsgakcaelqelglpvredvpligftrldyqkgldlkmaipelm 480
QY 481 REDVQFVMLGSGDPIFEGNMRSTESSYKDKFRGWGVFSPVSHRITAGCDILLMPSRFEP 540
Db 481 redvqfvmllsgdplifegnmrstessykdfrgwgvfsvpshritagcdillmpsrfeep 540
QY 541 CGLNQLYAMQYCRVPVYVHGTGGLRDVTFNPFKAGEEGTGWAFSLPTVDMKMLALRTA 600
Db 541 cglnqlyamqycrvpvyvhtgtgglrdvtfnfpkageegtgwafslptvdkmlalrta 600
QY 601 MSTFRHFKPSWGLMKRGTMDKHTWDHAFSSRSRSGSPSWNPTSCRGLGRSKCESPS 660
Db 601 matfrehkpswglmkrgtmdkhtwdhafsrsrsgspswntptscrglgrskcesps 660

QY 661 ALKTSSSSFRGPEGYPCTLRCPATVESQCCALLWFAGSRTYDGCAAAATVATSGGRQLQFW 720
Db 661 alktsssfgrgpegypctllrcpatvesqccallwfagsrtydgcaaaavtasgrqlqfw 720
QY 721 GIRKGCAGWLTAKHHSDGSLSVRVTAETIRNQLVTL 756
Db 721 girkgcagwltakhhsgslsvrvtaetirnlvltl 756
RESULT 2
AAW23937
ID AAW23937 standard; Protein: 671 AA.
XX
AC AAW23937;
XX
DT 21-MAY-1998 (first entry)
XX
DE Wheat soluble starch synthase partial sequence.
XX
KW Starch synthase; wheat; transgenic plant.
XX
OS Triticum aestivum L. cv. Florida.
XX
PN WO9745545-A1.
XX
PD 04-DEC-1997.
XX
PF 28-MAY-1997; 97WO-EP02793.
XX
PR 11-SEP-1996; 96DE-1036917.
PR 29-MAY-1996; 96DE-1021588.
XX
PA (AGRE) HOECHST-SCHERING AGREVO GMBH.
XX
PI Block M, Loerz H, Luetticke S, Froberg C, Kosmann J;
PI Walter L;
XX
DR WPI; 1998-032652/03.
DR N-PSDB; AAV01527.
XX
PT Nucleic acid encoding starch synthase enzymes from wheat - for
PT transgenic plants that produce modified forms of starch, useful e.g.
PT in foods, or for production of packaging materials and disposable
PT goods
XX
PS Claim 1; Page 47-50; 71pp: English.
XX
CC This amino acid sequence comprises a near full-length sequence for
CC a soluble starch synthase of summer wheat (cv. Florida). It was
CC deduced from a cDNA clone (AAV01527) isolated from a 21-day caryopsis
CC cDNA library. A granule-bound starch synthase (see AAW23938) has
CC also been identified. Isolated nucleic acids encoding these
CC enzymes can be inserted into vectors for production of transgenic
CC plants, particularly starch-producing plants, specifically wheat.
CC Use of the isolated nucleic acids, or of antisense sequences, allows
CC starch metabolism to be regulated in transgenic plants.
CC Overexpression may result in improved crop yield, while modification
CC of starch in planta may eliminate the need for subsequent
CC chemical/physical modification. Plants with altered levels of the
CC various isoforms of starch synthase will produce starch of different
CC chain length, amylose/amylopectin ratio, degree of branching,
CC phosphate content, gelatinisation behaviour, granule size and shape,
CC viscosity etc. The starch produced by such plants is useful
CC particularly in foods (especially bakery goods or pasta) or to
CC produce packaging materials or disposable goods, as well as in any
CC other known use of starch.
XX
SQ Sequence 671 AA;

Query Match 89.3%; Score 3612; DB 19; Length 671;
Best Local Similarity 100.0%; Pred. No. 4.4e-301;

Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	86	TOPPLDAGVGLAPDLLLEGLAEIDSIIIVAAASQDSEIMDANEOPOAKVTRTSVFVT	145
Db	1	tcppldagvgelapdlllleglaedidsilvaaseqdeimdanecpqqakvtrtsivft	60
QY	146	GEAAPYAKSGGLDVGCSLPILAAARGHRVWVMPRYLNGSSDKNYAKALYTAKHKIPC	205
Db	61	geaapyaksgglgdvcgslpialaarghrvmvmprylmgssdknyakalytakhikipc	120
QY	206	FGGSHEVTFEHEYRDNVDFVFDHPSYHRPGSLYGNDFGAFGDNQRYTLLCYAAEAPL	265
Db	121	fggshevtffheyrdndvdfvfdhpsyhrpgslygndfagfdnqrfytllyaaeapl	180
QY	266	ILELGGYIYGQCMFVNDWHASLVPLVLAARYPGVYRDRSRTLVINLHAOGVEPAS	325
Db	181	ilelgyiygqcmfvndwhaslvplvllaarkyrgvyrdrsrtlvihnlahqgvepas	240
QY	326	TYPDGLPPEWYGALWVPEWARRHALDKGEAVNFKGAVVTADRVTVSQGSWEVTT	385
Db	241	typdglppewygalewfpewarrhaldkgeavnflkgavvtadrvtvsqgswevtt	300
QY	386	AREGQGLNELLSRKSRLVINGINDINWNPRTTKCLPHHSVDDLKGAKCKAELQKE	445
Db	301	aeggqglnellssrksrlvngidindwmpdtkclphhsyvdldsgkakckaelqke	360
QY	446	LGLPVREDVPLIGFIRLDYOKGIDLIKMAIPELMREDYQFVMLGSGDPIFEQWMBSTES	505
Db	361	lglpvredvpllfgirldyqgldikmaipeimredyqfvmigsgdplfegwmbstes	420
QY	506	SYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEPGCLNQLYAMOYGTVPVHGTGLRD	565
Db	421	sykdkfrgwgfsvpshritagcdillmpsrfepcglnglyamoygtvpvhgtglrd	480
QY	566	TYETNPFGAKGEGTGWAFSPLTVDKMLWALTAMSTPREHKPSWEGMLKRGMTKDHFW	625
Db	481	tvetnfpfakgeegtgwafspitvdkmlwalrtamstprehkpsweglmkrgmtkdhfw	540
QY	626	DHAPSTSRSSGSPWNTPTSCRRGLGRSKCESPSALKTSSSFRGPEGYPCTLRCPATV	685
Db	541	dhapstsrssgspwntptscrrglgrskcespsalktsssfrrgepypcclrcpatv	600
QY	686	ESQACLLWFAGSRTYDGCAAAAVTASGGRLQFNGIRKCAAGWLTAKHHSDGSLSVRV	745
Db	601	esqaccllwfagsrtydgcaaaavtasgrrlqfngirkgaagwltakhhsgslsvrv	660
QY	746	TAIRNQLVTL 756	
Db	661	tairnqlvltl 671	
RESULT 3			
AAY09004			
ID	AAY09004 standard; Protein; 647 AA.		
XX			
AC	AAY09004;		
DT			
XX	05-JUL-1999 (first entry)		
XX			
DE	Wheat starch soluble synthase I (SSS I) amino acid sequence.		
XX			
KW	Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;		
KW	starch branching enzyme; starch soluble synthase; debranching enzyme;		
KW	endosperm; wheat; barley; granule-bound synthase; glutenin; starch;		
KW	grain softness protein I; bacterial isoamylase; glycogen synthase;		
XX	WSBE I-D4 gene.		
OS	Triticum tauschii.		
XX			
PN	W0914314-A1.		
XX			
PD	25-MAR-1999.		
XX			

PF	11-SEP-1998;	98WO-AU00743.	
XX			
PR	20-MAR-1998;	98AU-0002509.	
PR	12-SEP-1997;	97AU-0009108.	
XX			
PA	(CSIR) COMMONWEALTH SCI & IND RES ORG.		
PA	(GOOD-) GOODMAN FIELDER LTD.		
PA	(LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.		
PA	(AUSU) UNIV AUSTRALIAN NAT.		
XX			
PI	Li Z, Morell M, Rahman S;		
XX			
DR	WPI; 1999-229525/19.		
DR	N-PSDB; AAX34651, AAX34652.		
XX			
PT	New isolated cereal plant enzyme genes used for, e.g. expression of		
PT	antisense sequences of granule bound synthase		
XX			
PS	Claim 13; Page 95-97; 171pp; English.		
XX			
CC	The invention relates to a novel enzyme of starch biosynthetic pathway		
CC	in a cereal plant, where the enzyme is selected from starch branching		
CC	enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching		
CC	enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or		
CC	SBE I of rice or maize. The methods and products can be used for		
CC	targeting expression specifically to the endosperm of the seeds of cereal		
CC	plants such as wheat or barley. They can be used for the expression of		
CC	e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low		
CC	mol. wt. glutenin, grain softness protein I, bacterial isoamylase,		
CC	bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They		
CC	can be used for modifying the characteristics of starch produced by a		
CC	plant. The present sequence represents the wheat SSS I protein sequence.		
XX			
SQ	Sequence 647 AA;		
	Query Match	82.8%;	Score 3348; DB 20; Length 647;
	Best Local Similarity	99.7%;	Pred. No. 1.9e-278;
	Matches	626;	Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY	1	MAATGVAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSRGPAARPAQ	60
Db	1	maatgvagclapsvrlradpataaracsvvrarlrllargryvaelsregpaarpaq	60
QY	61	QQLAPPLVPGFLAPPAPAPQAPPTQPLPDAGVGLAPDLLLEGLAEIDSIIIVAA	120
Db	61	qqlapplvpgflappapapqaptpplpdagvgelapdlllleglaeidsilvaas	120
QY	121	EQDSEIMDANEQPAQVTRTSIVFTGFAAPYAKSGGLDVGCSIPILAAARGHRVMVMP	180
Db	121	eqdseimdanecpqaqvtrtsivftgeaapyaksgglgdvcgslpialaarghrvmvmp	180
QY	181	RYLNGSSDKNYAKALYTAKHKIPCFGGSHEVTFEHEYRDNVDFVFDHPSYHRPGSLY	240
Db	181	rylngssdknyakalytgkhikpcfggshvtfefheyrdndvdfvfdhpsyhrpgsly	240
QY	241	DNFGAFGDNQFRYTLCLYACEAPLILELGGYLYGQCMFVNDWHASLVPLVLAARYP	300
Db	241	dnfgafgdnqfrytllycyaaceaplillelgylygqcmfvndwhaslvplvllaaryp	300
QY	301	YGYRDRSRSTLVINLHAOGVEPASTYPDLGLPPEWYGALWVPEWARRHALDKGEAVN	360
Db	301	ygyrdrsrstlvihnlahqglepastypdlglppeygalewfpewarrhaldkgeavn	360
QY	361	FLKGAVVTADRVTVSQGSWEVTTAEGGQGLNELLSRKSRLVINGINDINWNPRTD	420
Db	361	flkgavvtadrvtvsqgswevttaeaggqglnellssrksrlvngidindwnptd	420
QY	421	KCLPHHSYVDLDSGKAKCKAELQKELGPVREDVPLIGFIRLDYOKGIDLIKMAIPELM	480
Db	421	kcclphhsyvdldsgkakckaelqkelpvredvpligfirdyokgidlikmaipelm	480
QY	481	REDVQFVMI GSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDTLLMPSRFP	540

```

|||||
Db 481 redvqfvmlysgdpifegwmrscssykdfrgwgfsvpvshritagcdillmpsfef 540
QY 541 CGLNQLYAMQYGVVPHVHGTGGLRDTVETENPFCAGKEEGTGNAFSPLTVDKMLWALRTA 600
Db 541 cglinqlyamqygtvpvvhgtgglrdtvetfnpgakgeegtgwafapltvdkmlwalrta 600
QY 601 MSTFREHKPSWEGLMKRGWTKDHTWDHA 628
Db 601 mstfrehkpsweglmkrgmtkdhtwdha 628

RESULT 4
AAW38218
ID AAW38218 standard; Protein: 649 AA.
AC AAW38218;
DT 22-MAY-1998 (first entry)
XX Maize starch synthase type I.
DE Maize starch synthase type I.
XX Maize; starch synthase type I; starch.
XX Zea mays.
XX
XX PN DEL9619918-A1.
XX PD 20-NOV-1997.
XX PE 17-MAY-1996; 96DE-1019918.
XX PR 17-MAY-1996; 96DE-1019918.
XX (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
XX Froberg C, Kossmann J;
XX WPI: 1998-000821/01.
XX DR N-PSDB; AAT95785.
XX
XX DNA encoding maize starch synthase type I protein - for producing
XX transgenic plants.
XX Claim 1: Pages 16-20; 23pp; German.
XX
XX The present sequence is maize starch synthase type I, useful
XX in the production of starch. Starch can be used in various
XX conventional starch applications, e.g. starch hydrolysate products,
XX foods, papermaking, adhesives, textiles, building materials, soil
XX stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal
XX briquettes, ore and coal slurries, foundry casting, rubber, leather
XX and synthetic polymers. The enzyme produces a starch stated to have
XX different physicochemical properties, especially viscosity and
XX gelling properties, from wild type starch.
XX
XX SQ Sequence 649 AA;

Query Match 68.0%; Score 2749; DB 19; Length 649;
Best Local Similarity 82.5%; Pred. No 5e-227;
Matches 518; Conservative 34; Mismatches 62; Indels 14; Gaps 4;

QY 4 TGVGAGCLAPSRLRADPATAARASACVVRVRA---RLRLRARGRYVAELSRGPAARPAQ 60
Db : ||| || : || : || : || : || : || : || : || : || : || : || : ||
14 savgaacll-----laraawpaavgdtrprlrqrvlrrrcvaelsregpaprmp 65
QY 61 QQLAPPLVPGFLAPPPAPPAQSPAPTPPLPDAGVCELAPDLLEGAEDSIDSIIVAA 120
Db : ||| ||| || : || : || : || : || : || : || : || : || : || : ||
66 allappilvpffla-ppaeptgepaltpppvpdaglvgve--pgliaegsldntvvva 122
QY 121 EODSEIMDANEQPAQKATRSIVVTGEAAPYAKSGGLGVCGSLPITALAARGHRVWVMP 180
||| ||| : || : || : || : || : || : || : || : || : || : || : ||

```

```

Db 123 eqdseivvgkeqarakvtqnivfvtgeaspyaksgglgdcgslpvalaarghrvmvmp 182
QY 181 RYLNSSDKNYAKALYTAKHIPCFGGSHVETFFHEHYRONVDWVFVDHPSYHRPGSLYG 240
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 183 rylngtsdknyanafytekhiripcfgehevtffheyrdsvdwfvdhpsyrpgnlyg 242
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
QY 241 DNFGAFGDNQFRYTLICYAACAPLILELGGYIYGQCMFVYNDWHAASLVPVLLAAKYRP 300
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 243 dkfgafgdnqfrytllicyaaceaplvlelgyiygqcmfvvndwhaslvpvllaakyrp 302
QY 301 YGVYRDSRSTLVTHNLAHQGVPEASTYDGLPEPEWYGALEWVPEWARHDLKGEAVN 360
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 303 ygvyksrslvlhmlanggvpeastypdlglppewygailewvpewarhaldkgeavn 362
QY 361 FLKGAVVTADRIVTVSQGYSWEVTTAREGGQGLNELLSRKSVLNGIYNGIDINDWNP 420
Db 363 flkgavvtadrivtvskgyswevttaeggglnellssrksvlnvgidindwnpatd 422
QY 421 KCLPHYSVDDLSGKAKKAELOKELGLPVREDVPLIFIGRDLQYKGLIDILKMAIFELM 480
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 423 kclpchysvddlsgakckgalkelglprdpvplifigrldyqkglidliqilpdlm 482
QY 481 REDVQFVMLGSGDPIFEQWMRSTESSYKDKFRGMVGFSPVSHRITAGCDILLMP 540
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 483 redvqfvmlysgdpeledwmrstesifkdkfrgwgfsvpvshritagcdillmpsfef 542
QY 541 CGLNQLYAMQYGVVPHVHGTGGLRDTVETENPFCAGKEEGTGNAFSPLTVDKMLWALRTA 600
Db 543 cglinqlyamqygtvpvvhgtgglrdtvenfnpgengegtgwafaplttenmlwlrta 602
QY 601 MSTFREHKPSWEGLMKRGWTKDHTWDHA 628
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 603 istyrehkssweglmkrgmtkdhtwdha 630

RESULT 5
AAW51231
ID AAR51231 standard; Protein: 626 AA.
XX
XX AC AAR51231;
XX
XX DT 16-NOV-1994 (first entry)
XX
XX DE Soluble rice starch synthetic enzyme.
XX
XX KW Rice; starch synthetic enzyme; transit peptide; amyloplast;
XX transition; protoplast; expression.
XX OS Oryza sativa.
XX
XX FH Key Location/Qualifiers
FT Active-site 1..113 /label= transit_peptide
FT /note= "claim 3"
FT Protein 114..626 /label= mat_protein
FT /note= "claim 1"
XX
XX PN JP06070779-A.
XX
XX PD 15-MAR-1994.
XX
XX PF 07-JUL-1992; 92JP-0179947.
XX
XX PR 07-JUL-1992; 92JP-0179947.
XX
XX PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX
XX WPI: 1994-128678/16.
XX DR N-PSDB; AAQ45183.
XX
XX PT Soluble rice starch synthetic enzyme gene and transit peptide -
XX for the efficient transport of heterologous proteins to

```



```

XX CC The sequence is that of maize starch synthase from pEXS2.
CC CC It can be used in the production of a hybrid polypeptide
CC CC comprising a starch-encapsulating region (SER) fused
CC CC to a payload protein. The hybrid polypeptide can be used to make
CC CC modified starches comprising the payload protein, selected from,
CC CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
CC CC immunoglobulins, etc. The modified starch can also be used
CC CC to provide grain feeds enriched in amino acids. By encapsulating
CC CC the payload protein in starch, it is more resistant to
CC CC degradation by stomach acids.
XX SQ Sequence 539 AA;

Query Match 59.1%; Score 2390.5; DB 19; Length 539;
Best Local Similarity 88.4%; Pred. No. 2.4e-196;
Matches 444; Conservative 23; Mismatches 32; Indels 3; Gaps 2;

QY 95 VGEIA-PDLLL--EGIAEDSIDSIIVAASEQDEIMDANEQQAQVTRSIIVFTGEAAY 151
DB 2 vaelrsrdlglepegiaegsidntvvvaseqdeivvgkeqarakvtqsvftgeaspy 61
QY 152 AKSGGLDVCGLPIALAAARHRVWVMPRYLNGSSDKNTAKALYTAKHKIKPCFGGSHE 211
DB 62 aks99gldvcgslpvalaarghrvvmvmprylngtsdknyanafytekhiripcfgehe 121
QY 212 VTFEYRDNDVWVFDHPHSYHRPGSLYGDNFGAFGDNQFRYTLICVAAACEAPLILELGG 271
DB 122 vtfefeyrdsdvwfvdhpsyhrpgnlygdkfgafgdnqfrytllicyaaaceaplilelg 181
QY 272 YIYGONCMFVNDWHASLVPLVLAAYRPGYVYRDSRSTLVIHNLAHQGVPEASTYPDLG 331
DB 182 yiy9qncmfvndwhaslvplvllaayrpygykdsrsilvihnlahqgvpeastypdlg 241
QY 332 LPPEWYGALEWPEWARRHALDKGEAVNFKGAVVTADRIYVTSOGYSWEVTTAGGQG 391
DB 242 lppewgaalewfpewarrhaldkgeavnfkgaavvtadriyvtvsgyswewttaggqg 301
QY 392 LNELLSRKSVLNGVINGIDINDWNPSTDCKLPHYSVDDLSGKAKCAELQKELGLPVR 451
DB 302 lnellsrksvlnvgingidindwnpstdcklphysvddlsgekakcaelqkglpvr 361
QY 452 EDVPLIGFGRLDYQKIDLIKMAIPELMREDVQFVNLGSGDPIFCGWMRSTESSYKDXF 511
DB 362 pdvpllgfgrldyqkgidliqilpdlmredvqfvmnlsgdpfelcdwmrstesifkdkf 421
QY 512 RGWGFSPVSHRITAGCDITLLMPSRFEPCGGLNOLYAMOYGTVPVVGHTVETEN 571
DB 422 rgwgfspvshritagcditllmpsrfepcglnlyamoygtvpvvhgtvettten 481
QY 572 PFGAKGEGTGWAFSLTVDKM 593
DB 482 pfgeqegtgwafaplttenm 503

RESULT 9
AAR99540
ID AAR99540 standard; Protein; 459 AA.
XX
AC AAR99540;
XX
XX 31-OCT-1996 (first entry)
XX
XX Soluble starch synthase (Isoform B).
XX
XX Soluble starch synthase; transformation; transgenic plant; starch;
XX amylose; amylopectin.
XX
XX Solanum tuberosum.
XX
XX DE4441408-AL.
XX
```

```

PD 15-MAY-1996.
XX
XX 10-NOV-1994; 94DE-4441408.
XX
XX 10-NOV-1994; 94DE-4441408.
XX
XX (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
XX
XX Abel GJ, Kossmann J, Springer F;
XX
XX WPI; 1996-240218/25.
XX
XX N-PSDB; AAT32325.
XX
XX DNA encoding soluble starch synthase of potato - used to produce
XX transgenic plants with increased prodn. of starch or able to produce
XX modified starches
XX
XX Claim 11; Page 20-22; 32pp; German.
XX
XX DNA sequences from the potato encoding soluble starch synthase
XX having this 459 C-terminal amino acid sequence or the 677
XX C-terminal amino acid sequence given in AAR99539 can be used to
XX identify and isolate homologous sequences encoding soluble starch
XX synthase and enzymes with similar activities from plants or other
XX organisms; to transform prokaryotic or eukaryotic cells; to produce
XX transgenic plants which synthesize starch of altered structure or
XX in increased yield.
XX
XX Sequence 459 AA;

Query Match 42.5%; Score 1718.5; DB 17; Length 459;
Best Local Similarity 70.6%; Pred. No. 8.4e-139;
Matches 303; Conservative 62; Mismatches 61; Indels 3; Gaps 2;

QY 203 IPCFGSHEVTFEYRDNDVWVFDHPHSYHRPGSLYGDNFGAFGDNQFRYTLICVAAACE 262
DB 14 vhcfdgaqevafyheyragdvwfvdhssyrrpgdygafgdnqfrytllicshaace 73
QY 263 APLILELGGYIYGONCMFVNDWHASLVPLVLAAYRPGYVYRDSRSTLVIHNLAHQGYE 322
DB 74 aplvlpigftygekcflfandcnaalvpllaakyrpygykdsrsilvihnlahqgye 133
QY 323 PASTYPDLGPEWYGALEWPEWARRHALDKGEAVNFKGAVVTADRIYVTSOGYSWE 382
DB 134 pavtynnlglpqwygavewifptwarahaldtgecnvfkgaiaavadriltvsggyswe 193
QY 383 VTTAGGGGGLNELLSRKSVLNGVINGIDINDWNPSTDCKLPHYSVDDLS--GKAKCKA 440
DB 194 ittpgegyglhellssrqsvlntgindvndwnpstdehshysindlspgkvqckt 253
QY 441 ELQKELGLPVREDVPLIGFGRLDYQKIDLIKMAIPELMREDVQFVNLGSGDPIFCGWM 500
DB 254 dlqkelglpvrpdcpligfgrldyqkgvdiilsaipelmqndvqvmvlgsgekqdw 313
QY 501 RSTESSYKDKFRGWGFSPVSHRITAGCDITLLMPSRFEPCGGLNOLYAMOYGTVPVVGHT 560
DB 314 rhtenlfkdkfrawgfnvpsvshritagcditllmpsrfepcglnlyamrygtipivht 373
QY 561 GGLRDTVETFPFGAKGE--EGTGWAFSLTVDKMJMALRTAMSTPREHKPSWEGLMKRCM 619
DB 374 ggldrtvkdfnpyaqegkgwtgwtfspltseklftiklairytehkssweglmkrgm 433
QY 620 TKDHTWDHA 628
DB 434 grdyswena 442

RESULT 10
AAW56487
ID AAW56487 standard; Protein; 698 AA.
XX
XX AAW56487;
AC
```

```
XX 11-SEP-1998 (first entry)
DT
XX Zea mays soluble starch synthase IIB.
DE
XX
XX SER; starch-encapsulating region; fusion vector;
KW soluble starch synthase IIB; glucosyl transferase.
KW
XX Zea mays.
XX
XX W09814601-A1.
PN
XX 09-APR-1998.
XX
XX 30-SEP-1997; 97WO-US17555.
XX
XX 30-SEP-1996; 96US-0026855.
XX
XX (EXSE-) EXSEED GENETICS LLC.
XX
XX Guan H, Keeling P;
XX
XX WPI; 1998-240100/21.
DR
XX N-PSDB; AAV29755.
DR
XX
XX Hybrid polypeptide comprising starch-encapsulating region and
PT protein - useful for, e.g. producing protein(s) resistant to
PT degradation by stomach acids
XX
XX Example 2; Page 36-38; 156pp; English.
PS
XX
XX The sequence is that of soluble starch synthase IIB.
CC
XX It can be used in the production of a hybrid polypeptide
CC comprising a starch-encapsulating region (SER) fused
CC to a payload protein. The hybrid polypeptide can be used to make
CC modified starches comprising the payload protein, selected from,
CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
CC immunoglobulins, etc. The modified starch can also be used
CC to provide grain feeds enriched in amino acids. By encapsulating
CC the payload protein in starch, it is more resistant to
CC degradation by stomach acids.
XX
XX Sequence 698 AA;
SQ
Query Match 29.8%; Score 1205.5; DB 19; Length 698;
Best Local Similarity 43.0%; Pred. No. 1.7e-94;
Matches 283; Conservative 83; Mismatches 221; Indels 71; Gaps 17;
QY 12 APSVFLRADPA-----TAARASACVVRARLRRLRGRVVAELSREGPAARPAQQQLAPP 66
DB 56 aasvraaaapaggeseeaaakssss-sqagavvggstakavdsasppnpltsapkqsqaam 114
QY 67 L--VPGFLAPPPPAQSPAPTQPLP-----DAG-----VGLAPDLLLEGTAEDSI 112
DB 115 qngtsgsssaastaapvsgpkadhpeapvtkreidasavkpepagddarpvesigiaepv 173
QY 113 DSIIIVASEQDS-----EIMDANEQPOA-KVTRISIVFTGEAAPYAKSGGLGDVCGSLP 165
DB 174 dakadaapatdaaasapydrédneggplagnvmnmvvvvassecapfcctgglgdvvgalp 233
QY 166 TALAARHRVVMVPRYLNGSSDKNYAKALYTAHIIKIPCGGSHVTFEHEYRNDNDVW 225
DB 234 kalargrhrvmvlpvry-----geyaearldlgvrrrykvagqdvsevfhsyldgvdv 287
QY 226 FVDHPSY--HRPGSLGDNFGAFDGNQFRYTLICYAACBAPLILELGGVIYGO--NCMFYVN 283
DB 288 fveapfrrhnniyy---gerldilkmlfckaavevpyapcggtvygdgnlvfian 344
QY 284 DHASLVPVLLAARYPGVYRDSRSTLVTHNLHQGVPEASTYPDGLPPEWYGALEWV 343
DB 345 dwhtallpylkayrdrnglmgqarsvlihviahqgrgvpdddfvndlp----- 394
```

```
QY 344 FPEWARRHALDK-----GEAVNFKCAVVTADRIVTVSOGYSWEVTTAEGGGLNEL 395
DB 395 -----ehyldhfklydnigghsnvfaaglkktadrvvtvsngmwlktseggwglhdi 448
QY 396 LSSRSKSVLNGIYNGIDINDWNPPTDKCLPH---HYSVDL--SGRAKCKAELQKELGLPV 450
DB 449 inqndwkllgglvngldmwnpavdhlsddaytnytfctldtgrkqckaaiqrqlglqv 508
QY 451 REDVPLIGTIGRLDYQKIGIDLKMAIPELMREDVQFVMIJGSDPIFEGWMRSTESSYKDK 510
DB 509 rddvpligfgrldhckgvdladaihwlagdgvqlvmlgtgradledmlrrfesehsdk 568
QY 511 FRGWGFSVPVSHRITAGCDILLMPSPFPCGLNQLYAMQYGTVPVVGCTGGRLDHFVTF 570
DB 569 vrawvgfsvplahrifitagadillmpsfepcglinqiyamaygtvpvvhavggldrcvaf 628
QY 571 NPEGAKGEGTGWAFSPLTVDKMLWALRTAMSTFRHKKPWSWGLMKRGMTKDHTWDHA 628
DB 629 dpf---ndtglgtwfdraeanrmidalschlttyrnykeswraetrargmaedlswdha 683
RESULT 11
AAW70892
ID AAW70892 standard; Protein; 804 AA.
XX
AC AAW70892;
XX
DT 23-AUG-1999 (first entry)
XX
DE Maize starch soluble synthase IIB.
XX
KW Non-glycogen-like polysaccharide production; fermentation;
KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;
KW non-starch branching gene; amylopectin; amylose; plant-like starch;
KW maize starch soluble synthase IIB.
XX
OS Zea mays.
XX
PN W09844780-A1.
XX
PD 15-OCT-1998.
XX
PF 03-APR-1998; 98WO-US06660.
XX
PR 04-APR-1997; 97US-0042939.
XX
PA (EXSE-) EXSEED GENETICS LLC.
XX
PI Guan H, Keeling PL;
XX
XX WPI; 1998-568285/48.
DR N-PSDB; AAV70938.
XX
XX Producing non-glycogen-like polysaccharides in bacteria, fungi or
PT plants - transformed with genes for enzymes involved in starch or
PT glycogen synthesis allows fermentative production of starches with
PT engineered properties
XX
XX Disclosure; Fig 50; 150pp; English.
XX
XX The specification describes a method for the production of
CC non-glycogen-like polysaccharides in a host. The method comprises
CC transforming a host, suitable for fermentation, with genes encoding
CC starch- or glycogen-synthesis enzymes, and fermenting the transformants.
CC The specification also describes hosts transformed with a gene active
CC in glycogen synthesis and at least one non-starch branching gene,
CC involved in production of amylopectin or amylose in its original host.
CC The method is used to produce plant-like starches by fermentation and
CC new starches in plants. These starches are useful for all food and
CC non-food applications of starch. The present sequence is used in
CC the course of the invention.
XX
XX Sequence 804 AA;
SQ
```


Query Match 29.8%; Score 1203.5; DB 19; Length 804;
 Best Local Similarity 41.7%; Pred. No. 3.2e-94;
 Matches 287; Conservative 85; Mismatches 229; Indels 87; Gaps 18;

QY 12 APSVRLRADPA-----TAARASACVVRRLRLARGRYVAELSRGPAARPAQOQLAPP 66
 Db 56 aasvraaapagseaeaksss-sqagavqgstakavdsasppnltspakgsaam 114
 QY 67 L--VPGFLAPPAPPAQSPATQPLP-----DAG-----VGLPADLLLEGIAEDSI 112
 Db 115 qngtsgsesaastapagpkadhpsapvtkreidasavkpepadgdarpviesigiae-pv 173
 QY 113 DSIIVAAASEQDS-----EINDANEQQA-KVTRSIIVFTGEEAARYAKSGGLGDDVCGSLP 165
 Db 174 dakadaapatdaasapydrednepglagpnvmvvaasecapfctg9lgvavgalp 233
 QY 166 IALAARGHRVVMVPRYLNGSSDKNYAKALYTAHKIKIPCGGSHEVTFEYRDNDVW 225
 Db 234 kalarrghrmvmpvpry-----geyaearldgvrrrykvaggsdsevtfhsyldgvfv 287
 QY 226 FVDHPST-HRPGSLYGNFAGFNQPRYTLICVAACEAPLILELGGYIYGO-NCMFVNV 293
 Db 288 fveappfrhchniyg--geridilkrmilfckaavevpwypagpgvtyv9dgnlvfian 344
 QY 284 DWHASLVPLLAARYPRYGVYRDSRSTLVTHNLAHQVEPASTYPDGLPPEWYGALEWV 343
 Db 345 dwhcallpvykayrdnglmqyarsvlvlniahggrgvddfvnfdlp----- 394
 QY 344 FPEWARRHALDK-----GEAVNFKGAVVTADRTVTVSQYSWEVTTAEGGQGLNEL 395
 Db 395 -----ehyldhfklydnigghsnvfaagktadrvtvtsngymwelktseggwghdi 448
 QY 396 LSSRSKSLVINGIDINDWNTDCKLPH-----HYSVDDL-SGAKCKAELOKELGLPV 450
 Db 449 inqndwklqivngidmsepnpavdvhllshddytntyfctldgkrqckaqlrglglv 508
 QY 451 REDVPLIGFTRLDYKGDILKMAPELMDREDVQFVMLGSGDPIFEGBMRSTESSYKOK 510
 Db 509 rddvpligfgrldhqdqdiadahlwtagdqvqlvmlgtgradledmlrrfesehsdk 568
 QY 511 FRGWGFSVPVSHRITAGCDILLMPSEPCGLNQLYAMQYGTVPVVGHTGLRDTVETF 570
 Db 569 vrawvfvplahritagadillmpsfepcgnlqlyamaygtvvpvhavvgglrdtvpaf 628
 QY 571 NPGKAGEECTGWAFSPLTVDMKMLARTAMSTFREHKPSWEGMLKRGMTKDHTWDHAP- 629
 Db 629 dpf---ndtglwtfdraeanrmidalschltttrnykeswcracrgmaedlsdwhaav 685
 QY 630 -----SSTRSSSGPSWT 642
 Db 686 lyedvlvkakyqvanlatrrrcrtwt 713

RESULT 12
 AAB37567
 ID AAB37567 standard; Protein; 799 AA.
 XX AAB37567;
 AC AAB37567;
 DT 01-MAR-2001 (first entry)
 DE Wheat starch synthase II SEQ ID NO: 4.
 KW Wheat; starch synthase; SSII; starch content; starch synthesis;
 KW food product; adhesive.
 XX Triticum aestivum.
 OS WO200066745-A1.
 PN 09-NOV-2000.
 XX PD

XX PF 28-APR-2000; 2000WO-AU00385.
 XX PR 29-APR-1999; 99AU-0000052.
 XX XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 XX Morell M, Li Z, Rahman S, Appels R;
 XX WPI: 2000-647602/62.
 DR N-PSDB; AAC86411.
 XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 PT WST-II, useful in modifying plant starch content and/or composition -
 XX Claim 19; Page 158-161; 21pp; English.
 XX The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building
 CC materials and packaging materials.
 XX Sequence 799 AA;
 SQ

Query Match 29.0%; Score 1174.5; DB 21; Length 799;
 Best Local Similarity 37.1%; Pred. No. 9.8e-92;
 Matches 286; Conservative 91; Mismatches 223; Indels 171; Gaps 19;

QY 7 GAGCLA-----PSVRLRADPATAARA-----SACVVRRLRLARGRYVAE---- 47
 Db 36 gagrlhwpwppqtrtdaggvaaraagkkdarvdddaasargprarrggaatkvaerdp 95
 QY 48 ----LSRE-----GPAARPAQOQLAPPPLVPGFLAPP-----PAPAQS 82
 Db 96 vktldraaegapappprqdaarp--psmngtpvngenkstgggatkdsghlpapara 153
 QY 83 PAP-TQPLPADAGVGEL-----APDLLLEGIAEDSIDSIIVAAASEODSET----- 126
 Db 154 phpstqnrpvngenkvanvasptsaieavpadsaatisiskapesvvpaeekpppsgs 213
 QY 127 -----MDANEQPAKV----- 137
 Db 214 nfvsasaprlididsdvepelkkgaviveeapnkalappaavqgdldwfkkyiffee 273
 QY 138 -----TRSIIVFTGEEAARYAKSGGLGDDVCGSL 164
 Db 274 pveakddgwavaddagsfehhqnhdsgplagenvmvmvvaecspwcktg9lgdvagal 333
 QY 165 PIALAARGHRVVMVPRYLNGSSDKNYAKALYTAHKIKIPCGGSHEVTFEYRDNDVW 224
 Db 334 pkalakrghrmvmpvpry---gdyeeaydvgrkyk--aagqdmvnyfhayldgvdf 387
 QY 225 VFVDHPST-HRPGSLYGNFAGFNQPRYTLICVAACEAPLILELGGYIYGO-NCMFV 282
 Db 388 vldaplfrrhrqediyg---gsrqeikrmilfckaavevpwypvpgvgdgnlvfia 444
 QY 283 NDWHASLVPLLAARYPRYGVYRDSRSTLVTHNLAHQVEPASTYPDGLPPEWYGALEW 342
 Db 445 ndwhcallpvykayrdnglmqytrslmvliahggrgvdpdefpfelpehyl----- 499
 QY 343 VPFEWARRHALDKGEAVNFKGAVVTADRTVTVSQYSWEVTTAEGGQGLNELSSRSKV 402
 Db 500 ---ehfrlydpvggehanyfaagkmdqvvvpspgylwelktveggwghdiirqndwk 556
 QY 403 LNGVINGIDINDWNTDCKLPH-----HYSVDDL-SGAKCKAELOKELGLPVREDVPLI 457
 Db 557 trgivngidmewnpdvdhksdgytnfsigtldsgkrqckealgrgelglqvradvp11 616

QY	458	GFTGRLDYQKIDLIKMAIPELMREDVQFVMLGSDPIFEGWNRSTESSYKDKFRGWGCF	517
Db	617	gfgrldgqkgveiadampwivsgdqvlgvmlgtgrhdlesmlrhferehdkvrgwvdf	676
QY	518	SVPSVSHRITAGCDILLMPSEPCGLNOLYAMOYGTVPVVHGTGGURDTVETFNPFPGKAG	577
Db	677	svrlahritagadallmpsfecpgcnqlyamaygtvpvvhavggvrdtvpfpdpl---	733
QY	578	EEGTGNAFSLPTVDKMLMALRTAMSTFREHKPSWEGLMKRGMTKDHWTWHA	628
Db	734	hsglgwtrfdraeahkklleaighclrtvdykeswrglqergmsqdfswaha	784
RESULT	13		
AAB37597			
ID	AAB37597	standard; Protein; 799 AA.	
XX			
AC	AAB37597;		
DT	01-MAR-2001	(first entry)	
DE		Wheat starch synthase II protein.	
XX			
XX		Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;	
KW		food product; adhesive.	
XX			
OS		Triticum aestivum.	
XX			
PN		WO200066745-A1.	
XX			
PD		09-NOV-2000.	
XX			
PF		28-APR-2000; 2000WO-AU00385.	
XX			
PR		29-APR-1999; 99AU-0000052.	
XX		(CSIR) COMMONWEALTH SCI & IND RES ORG.	
PA		(GOOD-) GOODMAN FIELDER LTD.	
PA		(LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.	
XX		Morell M, Li Z, Rahman S, Appels R;	
PI		WPI; 2000-647602/62.	
DR			
XX		Nucleic acid molecules encoding wheat starch synthase (WST)-I and	
PT		WST-II, useful in modifying plant starch content and/or composition -	
XX		Example 9; Fig 3; 21pp; English.	
PS		The present invention relates to novel protein and coding sequences from	
XX		wheat. The proteins are wheat starch synthases, designated SSII and	
CC		SSIII. These can be used in the modification of plant starch content or	
CC		composition, and to screen plants to identify mutations which affect	
CC		starch content and composition. The starch can then be used in food	
CC		products, such as flour, and in films, coatings, adhesives, building	
XX		materials and packaging materials.	
XX		Sequence 799 AA;	
SQ			
Query Match	29.0%;	Score 1174.5; DB 21; Length 799;	
Best Local Similarity	37.1%;	Pred. No. 9.8e-92;	
Matches 286;	Conservative 91;	Mismatches 223; Indels 171; Gaps 19;	
QY	7	GAGCLA---PSVRLRADPATARA-----SACVVRARLRLRARGRYVAE----	47
Db	36	gagrlhwppwppqrtdggvaaragkdkardvdddaasqrprarrgaatkvaerdp	95
QY	48	---LSRE---GPAARPAQOOLAPPLVPGFLAPPP-----PAPAQS	82
Db	96	vtldrdaaeggappaprgdaarp--psmngtpvngenkstgggatkdgslpapara	153

[illegible]

QY 48 ---LSRE-----GPAARPAQQOQLAPPLVPGFLAPP-----PAPA 80
Db 96 vktldraaeggpsppaarppsmngm-----pvngenkstgggatkdsigtpa 151
QY 81 QSPAP-TQPPLPDAGVCEL-----APDLLLGIAEDSDSIIV----- 117
Db 152 raphbstqnrapvngenkanvaspptsiaaaaaasdaatisidskapesvvpaehtpss 211
QY 118 -----AASQDSEIMDANEQ-----PQAKV----- 137
Db 212 gsnfessaaagdsdsvdeqelkgavvveeapkpalsppaapavqedlwfkkkyigf 271
QY 138 -----TRSIYFVTGEAAPYAKSGGLGDCVG 162
Db 272 eepveakddgravadagdsfehghndsgplagenvmnvvvaecspwcktgglgdvag 331
QY 163 SLPIALAAARGHRVVMVPRYLNGSSDKNAKALYTAKHIPCFGGSGHEVTFEYRDNV 222
Db 332 alpkaakagrhrvmvvprry-----gdyeeaydvgrkyk--aaqqdmevnyfhayidgv 385
QY 223 DWYFVDHPSY-HRPGSLYGNFAGCDNQFRYTLCLCYACEAPLIILELGGYIYQ-NCMF 280
Db 386 dfvfidaplfrhrqediyg--gsrqeimkrmilfckaavevpwhvpcggvpygdgnlvf 442
QY 281 VVNDWHASLVPLLAARYRPGYVYRDSRSTLVITHLAHQGVBPASTYPDLGLPPEWYAL 340
Db 443 iandwhtallpylkyayrhdgimqytrsimvihniahqgrgvpvdefpftelpehyl--- 499
QY 341 EWYFPEWARRALDKGEAVNFKGAVVTADRIVTVSQYSWEVTTAEGGGGLNELSSRK 400
Db 500 -----ehfrlydpvggehanyfaaglkmadqvvvvpspgylweiktvegglhdiirqnd 554
QY 401 SVLNGIUNGIDINDWNPPTDKCLPH----HYSVDL-SGKAKCAELQKELGLPVRDVP 455
Db 555 wktgrivngidnmewnpvhdahiksdgynfsirtidsgrkqckealqrelglqvradvp 614
QY 456 LIGFIRGLDYQKIDLIKMAIPELMRDYOVMLGSDPIFEGWMRSTESSYKDKPRGWV 515
Db 615 llgfrldgqkgveliadampwivsgdqlvmlgtgrhlesmlqhferrehdkvrgwv 674
QY 516 GFSVPVSHRITAGCDILLMPSPFEPGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGA 575
Db 675 gfsvriahritagadalmpsrfepeglnqlyamaygtvpvhavggldrtvppfdpf-- 732
QY 576 KGEEGTGWAFSPLTVDKMLMALMTAMSTFREHKPSWEGMLKRGMTKDHWTWDA 628
Db 733 -nhsglgwtfdraeahkliealghcirtyrdfkeswralqergmsqdsfweha 784

Search completed: March 28, 2002, 16:19:28
Job time: 5033 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 16:19:46 ; Search time 78.51 Seconds
(without alignments)
669.358 Million cell updates/sec

Title: US-09-674-824-2
Perfect score: 4044
Sequence: 1 MAATGVGAGCLAPSVRLRAD.....SDGSLSVRTAEIRNQLVTL 756

Scoring table: BLASUM62

Gapop 10.0 , Gapext 0.5

Searched: 335014 seqs, 69512255 residues

Total number of hits satisfying chosen parameters: 335014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New:*

1: /cgn2.6/ptodata/2/paa/US06_NEW_COMB.pep:*

2: /cgn2.6/ptodata/2/paa/US06_NEW_COMB.pep:*

3: /cgn2.6/ptodata/2/paa/US07_NEW_COMB.pep:*

4: /cgn2.6/ptodata/2/paa/US08_NEW_COMB.pep:*

5: /cgn2.6/ptodata/2/paa/US09_NEW_COMB.pep:*

6: /cgn2.6/ptodata/2/paa/US09_NEW_COMB.pep:*

7: /cgn2.6/ptodata/2/paa/US10_NEW_COMB.pep:*

8: /cgn2.6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4044	100.0	756	5	US-09-674-824-2
2	3612	89.3	671	6	US-09-952-677-2
3	3348	82.8	647	6	US-09-508-377-14
4	2749	68.0	649	6	US-09-931-297-2
5	1196.5	29.6	801	7	US-10-044-543-26
6	1174.5	29.0	799	7	US-10-018-418-4
7	1165	28.8	798	7	US-10-018-418-2
8	1148.5	28.4	799	6	US-09-952-677-6
9	1142.5	28.3	690	7	US-10-044-543-6
10	1120.5	27.7	597	7	US-10-018-418-6
11	875.5	21.6	616	6	US-10-044-543-14
12	863	21.3	564	6	US-09-595-328C-275
13	863	21.3	611	6	US-09-595-328C-274
14	862.5	21.3	528	6	US-09-595-328C-276
15	860	21.3	614	7	US-10-044-543-18
16	852.5	21.1	600	7	US-10-044-543-22
17	828.5	20.5	615	7	US-10-044-543-2
18	590.5	14.6	961	6	US-09-708-427-12954
19	590.5	14.6	1027	6	US-09-708-427-12953
20	590.5	14.6	1071	6	US-09-708-427-12952
21	568	14.0	533	7	US-10-044-543-10
22	535	13.2	867	6	US-09-708-427-7513
23	535	13.2	900	6	US-09-708-427-7512
24	535	13.2	1025	6	US-09-708-427-7511
25	496.5	12.3	1059	7	US-10-018-418-10

26 495.5 12.3 1628 7 US-10-018-418-8 Sequence 8, Appli

27 419 10.4 474 7 US-10-007-693-99 Sequence 99, Appl

28 389.5 9.6 476 7 US-10-007-693-69 Sequence 29, Appl

29 138 3.4 33 6 US-09-118-990A-29 Sequence 69, Appl

30 132.5 3.3 443 6 US-09-708-427-50399 Sequence 29, Appl

31 131.5 3.3 733 6 US-09-708-427-12776 Sequence 12776, A

32 131.5 3.3 756 6 US-09-708-427-10650 Sequence 10650, A

33 131.5 3.3 2110 6 US-09-614-150-1023 Sequence 1023, Ap

34 130.5 3.2 515 6 US-09-620-394B-6751 Sequence 6751, Ap

35 129.5 3.2 523 6 US-09-620-394B-6750 Sequence 6750, Ap

36 129.5 3.2 525 6 US-09-620-394B-6749 Sequence 6749, Ap

37 129.5 3.2 632 6 US-09-708-427-12777 Sequence 12777, A

38 127 3.1 695 6 US-09-708-427-10651 Sequence 10651, A

39 127 3.1 732 6 US-09-708-427-11216 Sequence 11216, A

40 125.5 3.1 746 6 US-09-708-427-11215 Sequence 11215, A

41 125.5 3.1 27 7 US-10-018-418-53 Sequence 53, Appl

42 123 3.0 358 6 US-09-897-516-7504 Sequence 7504, Ap

43 123 3.0 191 6 US-09-924-358-22 Sequence 22, Appl

44 122 3.0 194 6 US-09-708-427-5644 Sequence 56444, A

45 120.5 3.0

ALIGNMENTS

RESULT 1

US-09-674-824-2

; Sequence 2, Application US/09674824

; GENERAL INFORMATION:

; APPLICANT: Lorz, et al

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH CODE FOR ENZYMES DERIVED FROM WHE

; FILE REFERENCE: 1998/M 205/ FLH514413-3848

; CURRENT APPLICATION NUMBER: US/09/674,824

; PRIOR FILING DATE: 2000-02-16

; PRIOR APPLICATION NUMBER: PCT/EP99/03156

; PRIOR FILING DATE: 1999-05-07

; PRIOR APPLICATION NUMBER: DE 198 20607.0

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2

; LENGTH: 756

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-09-674-824-2

Query Match 100.0%; Score 4044; DB 5; Length 756;
Best Local Similarity 100.0%; Pred. No. 7.7e-317; Indels 0; Gaps 0;
Matches 756; Conservative 0; Mismatches 0;

Qy 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRRLRLRGRYVAELSRGPAARPAQ 60
Db 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRRLRLRGRYVAELSRGPAARPAQ 60
Qy 61 QOLAPPLVPGFLAPPPAPAPASPTOPPLPDAGVGLAPDLLLEGTAEDSIDSIIVAA 120
Db 61 QOLAPPLVPGFLAPPPAPAPASPTOPPLPDAGVGLAPDLLLEGTAEDSIDSIIVAA 120
Qy 121 EODSEIMDANEQOKVTRTSIVFTVGEAAPYAKSGGLGDCVCSLPALAAARGHRVWVMP 180
Db 121 EODSEIMDANEQOKVTRTSIVFTVGEAAPYAKSGGLGDCVCSLPALAAARGHRVWVMP 180
Qy 181 RYLNGSSDKNYAKALYAKHIKIPCFGSGSHEVTFEYHRDNVDMVFDHPSYHRPGSLYG 240
Db 181 RYLNGSSDKNYAKALYAKHIKIPCFGSGSHEVTFEYHRDNVDMVFDHPSYHRPGSLYG 240
Qy 241 DNFCAFQDNQFRYTLCLYACAPLILELGGYIYQNCMFVNDWHDHSLVPLLAAYRP 300
Db 241 DNFCAFQDNQFRYTLCLYACAPLILELGGYIYQNCMFVNDWHDHSLVPLLAAYRP 300
Qy 301 YGVTDRSRSTLVIHNLAHQGVPEASTYDGLPPEWYGALEWFPPEWARRHALDKGEAVN 360

|||||
Db 301 YGYRDSRSLVHNLAKHQVEPASTYDPLGLPPEWYGALEWVPEWARRHALDKGEAVN 360
QY 361 FLKGAVVADRIYTVSOGYSWEVTTAEGGGLNELLSRKS SVLNGVINGIDINDWNPSTD 420
Db 361 FLKGAVVADRIYTVSOGYSWEVTTAEGGGLNELLSRKS SVLNGVINGIDINDWNPSTD 420
QY 421 KCLPHYSVDLSGKAKAELOKELGVPVREDVPLIGFGRDYOYKGDILIKMAIPELM 480
Db 421 KCLPHYSVDLSGKAKAELOKELGVPVREDVPLIGFGRDYOYKGDILIKMAIPELM 480
QY 481 REDVQFVNLGSGDPIFEGMWRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEP 540
Db 481 REDVQFVNLGSGDPIFEGMWRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEP 540
QY 541 CGLNQLYAMQYGVTVVHGTGGLRDVETENPEFCAKEEGTGAFSPLTVDKMLWALRTA 600
Db 541 CGLNQLYAMQYGVTVVHGTGGLRDVETENPEFCAKEEGTGAFSPLTVDKMLWALRTA 600
QY 601 MSTFREHKPSWEGLMKGMTKDHTWDHAPSTSRSSGSPSWTNPSTCRRGLGRSKCESPS 660
Db 601 MSTFREHKPSWEGLMKGMTKDHTWDHAPSTSRSSGSPSWTNPSTCRRGLGRSKCESPS 660
QY 661 ALKTSSSFRGPEGYPTLRCPATVESQACLLWFAGRSRTYDGC AAAAVTASGGRQLQFW 720
Db 661 ALKTSSSFRGPEGYPTLRCPATVESQACLLWFAGRSRTYDGC AAAAVTASGGRQLQFW 720
QY 721 GIRKCAAGWLTAHHSDGSLSVRVTAIRNQLVTL 756
Db 721 GIRKCAAGWLTAHHSDGSLSVRVTAIRNQLVTL 756

RESULT 2
US-09-952-677-2
; Sequence 2, Application US/09952677
; GENERAL INFORMATION:

APPLICANT: Block, Martina

Lorz, Horst
Luticke, Stephanie
Walter, Lennart
Frohberg, Claus
Kossmann, Jens

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
FROM WHEAT WHICH ARE INVOLVED IN STARCH
SYNTHESIS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: James F. Haley, Jr., C/o Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/952,677

FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/196,390

FILING DATE: 19-Nov-1998

APPLICATION NUMBER: DE 196 21 588.9

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: DE 196 36 917.7

FILING DATE: 11-SEP-1996

APPLICATION NUMBER: PCT/EP97/02793

FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Haley, Jr., James F.

REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: AGREVO-9

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 596-9000

; TELEFAX: (212) 596-9090

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 671 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-952-677-2

Query Match 89.3%; Score 3612; DB 6; Length 671;
Best Local Similarity 100.0%; Pred. No. 3.7e-282;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TQPPLPDAGVGLAPDLLLEGTAEDSIDSIIVAASEQDSEINDANEQPOAKVTRISIVFT 145
Db 1 TQPPLPDAGVGLAPDLLLEGTAEDSIDSIIVAASEQDSEINDANEQPOAKVTRISIVFT 60
QY 146 GEAPYAKSGGLDVCGLPIALAAARHVRVMVMPRYLNGSSDKNYAKALYTAHKIKIP 205
Db 61 GEAPYAKSGGLDVCGLPIALAAARHVRVMVMPRYLNGSSDKNYAKALYTAHKIKIP 120
QY 206 FGSHEVTFEFHEYRONVDMVFDHPHSYHRPGSLYGNFAGFNQPRYTLCCYACEAPL 265
Db 121 FGSHEVTFEFHEYRONVDMVFDHPHSYHRPGSLYGNFAGFNQPRYTLCCYACEAPL 180
QY 266 ILEGGYIYQNCMVFVNDWHSALVPVLLAAKYRPGYVYRDRSRTLVIHNLHAHQVEPAS 325
Db 181 ILEGGYIYQNCMVFVNDWHSALVPVLLAAKYRPGYVYRDRSRTLVIHNLHAHQVEPAS 240
QY 326 TYPDLGLPPEWYGALEWVPEWARRHALDKGEAVNFKGAVVTADRIYTVSOGYSWEVTT 385
Db 241 TYPDLGLPPEWYGALEWVPEWARRHALDKGEAVNFKGAVVTADRIYTVSOGYSWEVTT 300
QY 386 AEGQGLNELLSRKS SVLNGVINGIDINDWNPSTDCLPHYSVDLSGKAKAELOKE 445
Db 301 AEGQGLNELLSRKS SVLNGVINGIDINDWNPSTDCLPHYSVDLSGKAKAELOKE 360
QY 446 LGLPVRDVPVLIQFGRDYOYKGDILIKMAIPELMREDVQFVNLGSGDPIFEGMWRSTES 505
Db 361 LGLPVRDVPVLIQFGRDYOYKGDILIKMAIPELMREDVQFVNLGSGDPIFEGMWRSTES 420
QY 506 SYKDKFRGWGFSVPVSHRITAGCDILLMPSPREPQGLNQLYAMQYGTVPVHGTGGLRD 565
Db 421 SYKDKFRGWGFSVPVSHRITAGCDILLMPSPREPQGLNQLYAMQYGTVPVHGTGGLRD 480
QY 566 TVETFNPFQAKGEEGTGWAESPLTVDKMLWALRTAMSTRFHKPSWEGLMKGMTKDHTW 625
Db 481 TVETFNPFQAKGEEGTGWAESPLTVDKMLWALRTAMSTRFHKPSWEGLMKGMTKDHTW 540
QY 626 DHAPSSTSRSSGSPSWTNPSTCRRGLGRSKCESPSALKTSSSFRGPEGYPTLRCPATV 685
Db 541 DHAPSSTSRSSGSPSWTNPSTCRRGLGRSKCESPSALKTSSSFRGPEGYPTLRCPATV 600
QY 686 ESQCACLLWFAGRSRTYDGC AAAAVTASGGRQLQFWGTRKCAAGWLTAHKIHS DGSLSVRV 745
Db 601 ESQCACLLWFAGRSRTYDGC AAAAVTASGGRQLQFWGTRKCAAGWLTAHKIHS DGSLSVRV 660
QY 746 TAEIRNQLVTL 756
Db 661 TAEIRNQLVTL 671

RESULT 3
US-09-508-377-14
; Sequence 14, Application US/09508377
; GENERAL INFORMATION:
; APPLICANT: KALEEN, ZHONGYILI
; APPLICANT: MORELL, MATTHEW

APPLICANT: RAHMAN, SADEOUR
TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS
FILE REFERENCE: 054270/0126
CURRENT APPLICATION NUMBER: US/09/508,377
CURRENT FILING DATE: 2000-06-09,
PRIOR APPLICATION NUMBER: AU PP 2509
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: PCT/AU98/00743
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: AU PP 9108
PRIOR FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 647
TYPE: PRT
ORGANISM: Triticum tauschii
US-09-508-377-14

Query Match 82.8%; Score 3348; DB 6; Length 647;
Best Local Similarity 99.7%; Pred. No. 6e-261;
Matches 626; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAATGVGAGCLAPSVLRADPATAARASACVVRRLRLARGRYVAELSRGPAARPAQ 60
Db 1 MAATGVGAGCLAPSVLRADPATAARASACVVRRLRLARGRYVAELSRGPAARPAQ 60
QY 61 QOLAPPLVPGFLAPPPAPAPQSPAPTOPPLPDAGVGLAPDLLEGIAEDSIDSIIVAS 120
Db 61 QOLAPPLVPGFLAPPPAPAPQSPAPTOPPLPDAGVGLAPDLLEGIAEDSIDSIIVAS 120
QY 121 EODSEIMDANEQOAKVTRISIVFTGEAAPYAKSGGLGDVCGSLPITALAARGHRVMVMP 180
Db 121 EODSEIMDANEQOAKVTRISIVFTGEAAPYAKSGGLGDVCGSLPITALAARGHRVMVMP 180
QY 181 RYLNGSSDKNYAKALYAKHITPCFGSGSHEVTFEYRDNDVDFVDPHPSYHRPGSLYG 240
Db 181 RYLNGSSDKNYAKALYAKHITPCFGSGSHEVTFEYRDNDVDFVDPHPSYHRPGSLYG 240
QY 241 DNFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLLAARYP 300
Db 241 DNFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLLAARYP 300
QY 301 YGYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPPEWYGALEWVPEWARRHALDKGEAVN 360
Db 301 YGYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPPEWYGALEWVPEWARRHALDKGEAVN 360
QY 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGGLNELLSSRKSVLNGIVNGIDINDNPTTD 420
Db 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGGLNELLSSRKSVLNGIVNGIDINDNPTTD 420
QY 421 KCLPHYSYDDLSGKAKCAEKLQELGLPVREDVPLIGF IGRLDYOKGIDLIKMAIPELM 480
Db 421 KCLPHYSYDDLSGKAKCAEKLQELGLPVREDVPLIGF IGRLDYOKGIDLIKMAIPELM 480
QY 481 REDVQVLMGSDPIFEGHMRSTESSYKDKFRGWGFSVPVSHRTAGCDILLMPSRREP 540
Db 481 REDVQVLMGSDPIFEGHMRSTESSYKDKFRGWGFSVPVSHRTAGCDILLMPSRREP 540
QY 541 CGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEGTGWAFSPLTVDKMLWALRTA 600
Db 541 CGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEGTGWAFSPLTVDKMLWALRTA 600
QY 601 MSTFREHKPSWGLMKRGTMDHTWDHA 628
Db 601 MSTFREHKPSWGLMKRGTMDHTWDHA 628

RESULT 4
US-09-931-297-2
; Sequence 2, Application US/09931297
; GENERAL INFORMATION:

APPLICANT: Jens Kossmann
Claus Froberg
TITLE OF INVENTION: Nucleic acid molecules encoding soluble
starch synthases from maize
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
CURRENT APPLICATION DATA: US/09/931,297
FILING DATE: 16-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/192,909
FILING DATE: <Unknown>
APPLICATION NUMBER: DE 196 19 918.2
FILING DATE: 17-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: GFB-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-96-9090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 649 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-931-297-2

Query Match 68.0%; Score 2749; DB 6; Length 649;
Best Local Similarity 82.5%; Pred. No. 9.4e-213;
Matches 518; Conservative 34; Mismatches 62; Indels 14; Gaps 4;
QY 4 TGVGAGCLAPSVLRADPATAARASACVVRRA---RLRLARGRYVAELSRGPAARPAQ 60
Db 14 SAVGAACLL-----LARAAPAAVGDRAAPRLQRLRRRCVAELSRGPAARMP 65
QY 61 QOLAPPLVPGFLAPPPAPAPQSPAPTOPPLPDAGVGLAPDLLEGIAEDSIDSIIVAS 120
Db 66 ALLAPPLVPGFLA-PPAETGEPALTPPPVDPAGLVGLVE--PEGIAEGSIDNTVVVAS 122
QY 121 EODSEIMDANEQOAKVTRISIVFTGEAAPYAKSGGLGDVCGSLPITALAARGHRVMVMP 180
Db 123 EODSEIVVKGEOAKVTONIVFTGEASPYAKSGGLGDVCGSLPITALAARGHRVMVMP 182
QY 181 RYLNGSSDKNYAKALYAKHITPCFGSGSHEVTFEYRDNDVDFVDPHPSYHRPGSLYG 240
Db 183 RYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFEYRDSDVDFVDPHPSYHRPGSLYG 242
QY 241 DNFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLLAARYP 300
Db 243 DKFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLLAARYP 302
QY 301 YGYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPPEWYGALEWVPEWARRHALDKGEAVN 360
Db 303 YGYKDSRSTLVTHNLAHQGVPEASTYPDLGLPPPEWYGALEWVPEWARRHALDKGEAVN 362
QY 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGGLNELLSSRKSVLNGIVNGIDINDNPTTD 420
Db 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGGLNELLSSRKSVLNGIVNGIDINDNPTTD 420

Db 363 FLKAVVATADRVTVTSKGYSEWTTAAGGGLNELLSSRKSVLNGVINGIDINDWNPATD 422
 Qy 421 KCLPHYSVDDLSGKAKKAELOKELGLPVREDVPLIGTGRDYOKGIDLIKMAIPELM 480
 Db 423 KCIPCHYSVDDLSGKAKKAELOKELGLPIRPDPVPLIGTGRDYOKGIDLIQIIPDLM 482
 Qy 481 REDVQFVMLGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILMPSRFP 540
 Db 483 REDVQFVMLGSGDPELEDWNRSTESIFKDKFRGWGFSVPVSHRITAGCDIILMPSRFP 542
 Qy 541 CGLNQLYAMQYGTVPVVGHTGGRLDRTVETFPNPGAKGEGTGWAFSLPTVDKMLWALRTA 600
 Db 543 CGLNQLYAMQYGTVPVVGHTGGRLDRTVETFPNPGAKGEGTGWAFSLPTVDKMLWALRTA 602
 Qy 601 MSTFREHKPSWEGMLKRGMTKDHWTWDA 628
 Db 603 ISTYREHKSSWEGMLKRGMTKDHWTWDA 630

RESULT 5
 US-10-044-543-26
 ; Sequence 26, Application US/10044543
 ; GENERAL INFORMATION:
 ; APPLICANT: Singletary, George
 ; APPLICANT: Zhou, Lan
 ; TITLE OF INVENTION: Novel Starch Synthese Polynucleotides
 ; TITLE OF INVENTION: and Their Use in the Production of New Starches
 ; FILE REFERENCE: 1144D
 ; CURRENT APPLICATION NUMBER: US/10/044,543
 ; CURRENT FILING DATE: 2002-01-11
 ; PRIOR APPLICATION NUMBER: 09/388,743
 ; PRIOR FILING DATE: 1999-09-02
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 801
 ; TYPE: PRT
 ; ORGANISM: Typha latifolia
 US-10-044-543-26

Query Match 29.6%, Score 1196.5; DB 7; Length 801;
 Best Local Similarity 42.6%; Pred. No. 1e-87;
 Matches 269; Conservative 80; Mismatches 200; Indels 83; Gaps 15;
 Qy 42 GRYVAELREGPAARPAQQOQLAP-----PLVPGFLAPPPP-----APAQSPAPT 86
 Db 109 GRYKS-----GAVPNYVSOAQDDTSENPLVNSFGSPKDNVEAVEFQVRSQAVDAFG 161
 Qy 87 QPPLPDAGVGLAPDLLLEG-----IAEDSDTSIIVAASEODSMDANEOP- 133
 Db 162 RPEESLGTITLSPFYLEAESDGAKEAEDLVEAKLDSVHV---KDDLNPGEENEVPL 218
 Qy 134 --QAKVTSISIVFTGEAAPAKSGGLGDCVCSLPTALAARGHRVVMVPRYLNGSSDKNY 191
 Db 219 PLAGANVMNIIIVAAECAPSKTGGLGDVAGALPKALARRHRVVMVAPRGNTAEPODI 278
 Qy 192 AKALYTAKHKIPCFGSGSHVETFFHEYRDNDVMDVFDVDRHPSY-HRPGSLYDGNFGAFGDNQ 250
 Db 279 GVRKYKVVH-----GQDMEVTFYHAYIDGVDFVMDSPDFRHRGNRIYECN---RVDIL 329
 Qy 251 FRYTLCLYCAACBAPLILEGGYIYGQ--NCMFVNDWNASLVPVLLAARYPGYVYRDSRS 309
 Db 330 KRMILFCKAAVEVPHVPCGGFCYDGNLAFITNDWHTALLPVLKAYIRONGMLMKYARS 389
 Qy 310 TLVIHNLAHQGVPEASTYDPLGLPPEWYGALEWVFPWARHALD-----KGEAYNF 361
 Db 390 VLVIIHIAHQGVPVDDDFKVLGLD-----HYLDLFRIDPVGGEHLNI 433
 Qy 362 LKGAVVATADRVTVTSKGYSEWTTAAGGGLNELLSSRKSVLNGVINGIDINDWNPATD 421
 Db 434 FAAGLKAADRVTVTVSHGYAWELKTSEGGWGLHEIINESNWKFGIYVNGIDAKESPEFDV 493

Qy 422 CLPH-----HYSVDDLS-CKAKKAELOKELGLPVREDVPLIGTGRDYOKGIDLIKMAI 476
 Db 434 HLKSDGYTYNSLDTLEMGKPVCKAALQREVGLPVHDNVPILIAFICRDLHQKGVDLIAEAM 553
 Qy 477 PELMEDREDVQFVMLGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILMPS 536
 Db 554 PWIVSHDVQVVMVLTGTRQDLENLLNFEGQHRDKYRANVAFSVKMAHRTITAGADILMPS 613
 Qy 537 REPCGLNQLYAMQYGTVPVVGHTGGRLDRTVETFPNPGAKGEGTGWAFSLPTVDKMLWA 596
 Db 614 REPCGLNQLYAMQYGTVPVVGHTGGRLDRTVETFPNPGAKGEGTGWAFSLPTVDKMLWA 670
 Qy 597 LRTANSTREHKPSWEGMLKRGMTKDHWTWDA 628
 Db 671 LNNCLNTYWNKYDSWKGGLQTRGMQDLSMDNA 702

RESULT 6
 US-10-018-418-4
 ; Sequence 4, Application US/10018418
 ; GENERAL INFORMATION:
 ; APPLICANT: Morell, Matthew
 ; APPLICANT: Li, Zhongyi
 ; APPLICANT: Rahman, Sadequr
 ; APPLICANT: Appels, Rudolph
 ; TITLE OF INVENTION: Genes Encoding Wheat Starch Synthases and Uses Therefor
 ; FILE REFERENCE: 127-01
 ; CURRENT APPLICATION NUMBER: US/10/018,418
 ; CURRENT FILING DATE: 2001-10-29
 ; PRIOR APPLICATION NUMBER: PCT AU00/00385
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: AU PQ0052/99
 ; PRIOR FILING DATE: 1999-04-29
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 799
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 US-10-018-418-4

Query Match 29.0%, Score 1174.5; DB 7; Length 799;
 Best Local Similarity 37.1%; Pred. No. 5.9e-86;
 Matches 286; Conservative 91; Mismatches 223; Indels 171; Gaps 19;
 Qy 7 GAGCLA-----PSVRLRADPATAARA-----SACVVRARLRRLRGRVYAE---- 47
 Db 36 GAGRLHWPWPQPQRTARDGGVAAAGKKDARVDDDAASARQPRARRGGAATKVAERRDP 95
 Qy 48 ---LSRE-----GPAARPAQQOQLAPPLVPGFLAPP-----PAPAGS 82
 Db 96 VKTLDRDAEAGGAPAPAPRQDAARP--PSMNGTPVNGENKSTGGGATKDSGLPAPARA 153
 Qy 83 PAP--TQPLPDAGVGL-----APDLLLEGTAEIDSIIVAASEQDSEI----- 126
 Db 154 PHFSTQNRVPGENKANVASPPTSIAEVVAPDSAAATISIDKAPESVVPAAEKPPSSGS 213
 Qy 127 -----MDANEQPOAKV----- 137
 Db 214 NFVVSASAPRLDIDSVEPELKKGAVIEBAPNPKALSPAPAVQBDLWDFKYYTGFE 273
 Qy 138 -----TRSVFVTGAAAPYAKSGGLGDCVCSL 164
 Db 274 PVEAKDDGWAADDAGSFEHHQNHDSGLAGENVMNVVVAACSPWCKTGGGLDVAGAL 333
 Qy 165 PIALAARGHRVVMVPRYLNGSSDKNYAKALYTAKHKIPCFGSGSHVETFFHEYRDNDV 224
 Db 334 PKALAKRGHRVVMVPRY----GDYEEAYDVGVRKYK--AAGQDMEVNTFHYAIDGVDF 387
 Qy 225 VFVDHPSY-HRPGSLYDGNFCAGFCDNQFRYTLCLYCAACEAPLILELGGYIYGQ--NCMFVV 282
 Db 388 VFIDAPLFRHQRQEDIYG----GSROEIMKRMILFCKAAVEVPHVPCGGVPGDGNLFIA 444

ATTORNEY/AGENT INFORMATION:
NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: AGREVO-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-952-677-6

Query Match 28.4%; Score 1148.5; DB 6; Length 799;
Best Local Similarity 37.0%; Pred. No. 7.2e-84;
Matches 286; Conservative 86; Mismatches 226; Indels 175; Gaps 20;

Qy 7 GAGCLA-----PSVRLRADPATAA-----RASACVVRARLRRLARGRYVAE----- 47
Db 36 GAGRLHPWPPTARDGVAALAAAGKDGADIDDAASVQPRALRGGAATKVAERRDP 95
Qy 48 ---LSRE-----GAARPAQOQLAPLPVPGFLAPP-----PAPA 80
Db 96 VKTLDRAAEGGSPPAARDQDARPPSMNGM-----PVNGENKSTGGGATKDSGLPTPA 151
Qy 81 QSPAP-TOPPLPDAGVGL-----APDLLLEGIAEDSIDSIIV----- 117
Db 152 RAPHESTNRPVNGENKANTASPTSIATAAASDAATISIDKAPESVVPAAEKTPSS 211
Qy 118 -----AASEQDSEINDANEQ-----POAKV----- 137
Db 212 GSNFESSASAPGSDTVSDVEQELKKGAVVVEEAPKPKALSPAPAPVQEDLWPKKYGIF 271
Qy 138 -----TRSIIVTGEAAPYAKSGGLGVCG 162
Db 272 BEPVENKODGRAVADAGSFEHQHNDGSLAGENVVNVVAAECPSPWCKTGGGLGVAG 331
Qy 163 SLPIALARGHRVVMVPRYLLNGSSDKNYAKALYAKHIKPCPGSGSEVTFEHEYRDNV 222
Db 332 ALPKALAKRGHRVMVWPY-----GDYBEADVGVKRYK-----AGQDMEVNFIHAYIDGV 385
Qy 223 DMVFDVDPHY-HRPGSLYDNGFAGDQFRYTLACYAAEAPLLELGGYIYGO-NCMF 280
Db 386 DFVFDAPLFRHQEDYIY-----GSRQEIIMKRMILFCKAAVEVPHVPCGVPYGDGNLVE 442
Qy 281 VVNDHSLVPLLAARYPYGVYRDSRSTLVIIHNLAHQGVPEPASTYDGLGLPPEWYCAL 340
Db 443 IANDWHTALLPVLYKAYYRDHLMQYTRSIMVIHIAHQGRPVDFEFPTELPEHYL----- 499
Qy 341 EWVPEWARHRLDKGEAVNFKLGAVVTADRTVTSQGSWEVTTAEGGQGLNELLSRK 400
Db 500 -----EHRFLYDPVGEHANYPAAGLKMDQVQVYSPGYLWELKTVEGSGWGLDRIQND 554
Qy 401 SVLNGLVNGIDINDNPTDKCLPH-----HYSVDDL-SGKAKCKAELQELGLPVREDVP 455
Db 555 WKTRGVNGIDNMENNPENVDHLKSDGYTNFSLRTLDGSKRQCKEALQRELGLQVRADVP 614
Qy 456 LIGFIRGLDYQKIDILKMAIPELMREDVQFVNLGSDPIFGGWRMSTESSYKDKFRGWV 515
Db 615 LLGFIRGLDGOKGVEIIADAMPVIVSQDQLVNLGTRHDLHESLMQGHFEREHHDKVRGWV 674
Qy 516 GFSVPSVSHRITAGCDILLMPSPFCGLNQLYAMQYGTVPVVGHTGDRDVTETNPFGA 575
Db 675 GFSVRLAHRITAGADALLMPSPFCGLNQLYAMAYGTVPVHAVGGLRDTVPPEDPF--- 732
Qy 576 KGEETGWAFTPLTVDKMLWALTAMSTFREHKPSWEGLMKRCMTKDHWDHA 628
Db 733 -NHSGLGWTDFRAEAHKLIEALGHCLRTYRDFKESWRAEQRMGSQDFSWEHA 784

RESULT 9
US-10-044-543-6
; Sequence 6, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthase Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches
; FILE REFERENCE: 1144D
; CURRENT APPLICATION NUMBER: US/10/044,543
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Curcuma zedoaria
US-10-044-543-6

Query Match 28.3%; Score 1142.5; DB 7; Length 690;
Best Local Similarity 47.5%; Pred. No. 1.8e-83;
Matches 238; Conservative 76; Mismatches 150; Indels 37; Gaps 11;

Qy 140 SIVFTVTEAAPYAKSGGLGVCGSLDIALAARGHRVVMVPRYLLNGSSDKNY--AKALYT 197
Db 200 NIILVAEAPWPKTGGGLGVGALPKALAKRGHRVMVSPRY-----GNYPEPEKIGN 253
Qy 198 AKHIKPCPGSGSEVTFEHEYRDNDVDFVDPHPSYHRPGS-LYGDNFAGFGONQFRYTL 256
Db 254 LKRYKVD--QGDMEIKYHYTIDSVDFIDSPIFRHIGNDIYGGN---RVDILKRMVLF 308
Qy 257 CYAACAPLLELGGYIYGO-NCMFVNVNDHSLVPLLAARYPYGVYRDSRSTLVIIH 315
Db 309 CKAAVEVPHVPCGFCYGDGNLVEIANDWHTSLLPVYLKACFRDRGLMTYARCLLVIIH 368
Qy 316 LAHQGVPEPASTYDGLPPEWYCALEWPEWARHRLDK-----GEAVNFKLGAVVTADRI 372
Db 369 IAHQGRGLDDFSYDLPDHDIDSR-----LDDPVGGEHFNIFAAGIRANDRV 417
Qy 373 VTSQGSWEVTTAEGGQGLNELLSRKSVLNGIDINDNPTDKCLPH-----HYS 428
Db 418 VTSYSHGAWELKTSEGOWGLHELINECHWKHFHGLVNGIDTHSNPKFADHLNSDGYTNET 477
Qy 429 VDDL-SGKAKCKAELQELGLPVREDVPLIGFTIGRLDYQKIDILKMAIPELMREDVQFV 487
Db 478 LETLEMGKACQKALQREFGLPVDRDDVPILAFIQLRDLHQKIDILAEAMHMLVGGDLQII 537
Qy 488 MLGSDPIFGGWRMSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSPFCGLNQLY 547
Db 538 MLGTRDLEDMRLRRFEREHRGKVRGWGFSVKMAHRTAGADALLMPSPFCGLNQLH 597
Qy 548 AMQYGTVPVVGHTGGLRDTVETNPFCAKEEGTGWAFSPLTVDKMLWALTAMSTFREH 607
Db 598 AMYGTIPVHAVGGLRDTVQQQDFPF---NETGLGWTDFRAEAHRMTVALGHCLNTYRN 654
Qy 608 KPSWEGLMKRCMTKDHWDHA 628
Db 655 KESWVGLQKRGMMQDLSWESA 675

RESULT 10
US-10-018-418-6
; Sequence 6, Application US/10018418
; GENERAL INFORMATION:
; APPLICANT: Morell, Matthew
; APPLICANT: Li, Zhongyi
; APPLICANT: Rahman, Sadequr
; APPLICANT: Appels, Rudolph
; TITLE OF INVENTION: Genes Encoding Wheat Starch Synthases and Uses Therefor

FILE REFERENCE: 127-01
CURRENT APPLICATION NUMBER: US/10/018,418
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: PCT AU00/00385
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: AU PQ0052/99
PRIOR FILING DATE: 1999-04-29
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 597
TYPE: PRT
ORGANISM: Triticum aestivum
US-10-018-418-6

Query Match 27.7%; Score 1120.5; DB 7; Length 597;
Best Local Similarity 41.7%; Pred. No. 8.6e-82;
Matches 254; Conservative 83; Mismatches 199; Indels 73; Gaps 16;
QY 37 RRLARGRYVAELSRGPAARPAQOQOOLAPPLVPG-----FLAPPPAPAPQSPAPT 86
Db 30 QELKKAVVVE---EAP-----KPKALSPAPAAVQEDLWDFKKYIGFPEPVEAK----- 76
QY 87 QPPLDAGVGEAPLAPLLLEGIAEDSDSIIVAASEQDSEIMDANEQOQAKVTRSIYFVVG 146
Db 77 -----DDG-----RAVADD-----AGSFEHQHSDGPLAGENV-MNVVVVAA 113
QY 147 EAPYAKSGGLGVCVCSLPALAAARGHRVVMVPRYLNSSDKNYAKALYAKHIKIPCF 206
Db 114 ECSPMCKTGLGVDVAGALPKALAKRGHRVMVVPYR---GDYEETDVGVRKYK--AA 167
QY 207 GGSHEVTFEYRDNDVDFVDPHSY--HRPGSLYGNFGAAGDNQPRYTLCCYAAACEAPL 265
Db 168 GQDMEVNYFAYIDGVDFVDFIDAPLFRHREDIYG---GSRQIMKRMILFCKAAVEVPW 224
QY 266 ILELGGYIYQ--NCMEVNDWHASLVPVLLAAKRYRPGYVYRDSRSLTVIHNLAHQGVPA 324
Db 225 HVPCCGVPGVGDGNLVFIANDWHFTALLPVYLKAYYRDHGLMQYTRSIMVHNIAHQGRGPV 284
QY 325 STYDPLGLPEWVGALEWVPEWARHRLDKGEAVNLFKAGVYADRIYTVVSOGYSWEVT 384
Db 285 DEFPFTPELPHYL-----EHFRLYDPVGGEHANYFAAGLKMADQVYVSPGYLWELK 336
QY 385 TAEQGGGLNELLSRKSVLNGVINGIDINDNPTTKCLPH----HYSVDDL--SGKAKCK 439
Db 337 TVEGGGLHDIIIRQNDWKTRGIVGIDNMEVNEVDALHLSKDYTNFSLRTLDSGRQCK 396
QY 440 AELQKELGPLVREDVPLIGFGRDLQKIDLIKMAIPELMDREDVQVMLGSGDPIFEGW 499
Db 397 EALQRELGLQVRADVPLLGLFGRLDQKQGEIITADAMPVISOVQLVMLGTGRHDLSE 456
QY 500 MRSTESSYKDKFGWGVFSVPVSHRITAGCDILLMPSRPEPCGGLNOLYAMQYGVTPVHG 559
Db 457 LQHFREHHDKVRGWGVFSVRLAHRITAGADALLMPSRPEPCGGLNOLYAMAYGVTPVHA 516
QY 560 TGLRDTVFENFPKAGEGCTGWAESPLTVDKMLALRTAMSTFREHKPSWEGLMKRG 619
Db 517 VGLRDTVPFPDFP-----NHSGLGWTFDRAEHLKLIHALGHLRTYRDFKESWALQERG 573
QY 620 TKDHTWDHA 628
Db 574 SQDFSWEHA 582

RESULT 11
US-10-044-543-14
Sequence 14, Application US/10044543
GENERAL INFORMATION:
APPLICANT: Singletary, George
APPLICANT: Zhou, Lan
TITLE OF INVENTION: Novel Starch Synthesis Polynucleotides
TITLE OF INVENTION: and Their Use in the Production of New Starches

FILE REFERENCE: 1144D
CURRENT APPLICATION NUMBER: US/10/044,543
CURRENT FILING DATE: 2002-01-11
PRIOR APPLICATION NUMBER: 09/388,743
PRIOR FILING DATE: 1999-09-02
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 616
TYPE: PRT
ORGANISM: Canina edulis
US-10-044-543-14
Query Match 21.6%; Score 875.5; DB 7; Length 616;
Best Local Similarity 38.8%; Pred. No. 4.6e-62;
Matches 199; Conservative 90; Mismatches 171; Indels 53; Gaps 13;
QY 140 SIVFVTGEAAPYAKSGGLGVCVCSLPALAAARGHRVVMVPRYLNSSDKNYAKALYAK 199
Db 91 NLVFGVCEVAPWSKTGGLDGLRGLPAPMAAIGHRVMTVVPYR-----DOYKDIWDTSV 144
QY 200 HKIPCFGGSHEVTFEYRDNDVDFVDPHSY-----HRPGSLYGDNFGA--EGDNQPR 252
Db 145 PVELKVGDKIETVRFPHFCYKRGVDRVDFVDPHMPLEKVMGKTGKLYGPVTGDIADNQLR 204
QY 253 YTLCCYAAACEAPLILLEG--YI---YQNCMEVNDWHASLVPVLLAAKRYRPGYVYRDS 307
Db 205 FSLCLLALEAPRLNLNNSKYSGPYGDDVFIANDWHASALLPCYLKTMVQSHGIYMNA 264
QY 308 RSTLVIHNLAHQGVPEASTYDPLGLPEWVGALEWVPEWARHRLD---KGEAVNFKG 364
Db 265 KVAFCIHNTAYOGRFAFSDFELLNLPNFKSSFDPM-----DGYDKPVKGRKINMKA 317
QY 365 AVTADRIYTVVSOGYSWE--VTTAEGGQGLNELLSRKSVLNGVINGIDINDNPTDKCL 423
Db 318 GIIEORCLTVSPYIAQELVSGVEKVELGNIL--BMKTCIGVINGMDTTEWNPPLTDKYI 375
QY 424 PHYSVDD--LSGKAKCKAELOKELGPLVREDVPLIGFGRDLQKIDLIKMAIPELMRE 482
Db 376 STNYDATTVDLAKPLCKEALQAEGLPVVKNKLVLAFAVRLDQKGSIDILAAAIPELCE 435
QY 483 DVQVFMVGLSGDPIFEGWMRSTESSYKDKFGWGVFSVPVSHRITAGCDILLMPSRPEPCG 542
Db 436 NQOVIVLGTGKKLESELTLLEEMFPDKFRAHLKFNVPVLAHIMAGADILVIPSREPCG 495
QY 543 LNOLYAMQYGVTPVVGHTGGLRDTVFENFPKAGEGCTGWAESPLTVD-----K 592
Db 496 LIQLQAMRYGTLPMCSTTGLVDTVK-----EGFTGFHMGPFVSVECDAYDKADVQK 546
QY 593 MLWALRTAMSTFREHKPSWEGLMKRGMTKDHTW 625
Db 547 IVETTRALKVY--GTPAFVEMIKMCMNODLSW 577

RESULT 12
US-09-595-328C-275
Sequence 275, Application US/09595328C
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Thereby
FILE REFERENCE: 2750-0946P
CURRENT APPLICATION NUMBER: US/09/595,328C
CURRENT FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 3107
SEQ ID NO 275
LENGTH: 564
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature

LOCATION: 1..564
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..564
OTHER INFORMATION: Ceres Seq. ID 1009070
US-09-595-328C-275

Query Match 21.3%; Score 863; DB 6; Length 564;
Best Local Similarity 38.1%; Pred. No. 4.1e-61;
Matches 203; Conservative 95; Mismatches 183; Indels 52; Gaps 15;
QY 119 ASEQDSSEIMDANEOPQAKVTRTSIVFTGAEAPYAKSGGLGVDGCSLPALAAARGHRVMV 178
Db 18 SSRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLGDLGGLPPALAAARGHRVMTI 76
QY 179 MPRYLSSGSKNAYKALYTAHKKIPCFGGSHEVTFEYRDNDVDFVDPHPSY-----H 233
Db 77 CPYI-----DQYKDAWDTCVVQIKVGDKNVRFHFCHYKRGVDRVDFVDPHIFLAKVVG 130
QY 234 RPSG-LYGDNFGA-FGDNQRYTLLCYAAACEAPLILELGG--YI---YGQNCMFVNDWH 286
Db 131 KTSKIKYGPITGVYNDNQLRFSLLCQAALAPQVNLNLSKYSFGPYGEDVVFVANDWH 190
QY 287 ASLVPVLLAAKYRPGYVYRDSRSTLVHNLAHQGVPEASTYPDLGLPPEWYGALEWVFE 346
Db 191 TALLPCYLKSMYQSRGYVMNAKVFCIHNIAVQGRFAFDYDYLNLPISEKSSFDF---- 246
QY 347 WARRHALD-----KGEAVNFKAGVYADRTVTSOGYSWE-VTTAEGQGLNELSSR 399
Db 247 -----MDGYEKPVKGRKINMKAAILEAHRLVTVSPYQAELISGVDRGVELHLYL--R 298
QY 400 KSVLNGIVNGIDINDNPTDKCLPHHYSVDLLS-GKAKCKAELOKELGLPVREDVPLIG 458
Db 299 MKTVSGIINGMDVQWENPSTDKVIDIKYDITVTDAKPLIKEALQAAVGLPVRDVPVIG 358
QY 459 FIGRLDYOKGIDILKMAIPELMREDVQVMLGSDPIFEGWMRSTESSYKDKFRGWGFS 518
Db 359 FIGRLEQKSDILVEAISKFMGLNVQVILGTGKKKMEAQILELEKFPKAVGVAKFN 418
QY 519 VPVSHRTACDILLMPSRFEPCGLNQLYAMQYGTVPVHGTGGLRDTVET-----FNP 572
Db 419 VPLAHMITAGADFIVPSRFEPCGLIQLHAMRYGTVPVASTGGLVDTVKDGYTGFIHR 478
QY 573 FGAKGEGTGWASPLTVDKMLWALRTAMSTFREHKPSWEGLMKRGMTKDHTW 625
Db 479 FNVKCE-----VVDPPDVIATAKAVTRAVAVY--GTSAMQEMVKNCMQDQFSW 524

RESULT 13
US-09-595-328C-274
Sequence 274, Application US/09595328C
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-0946P
CURRENT APPLICATION NUMBER: US/09/595,328C
NUMBER OF SEQ ID NOS: 3107
SEQ ID NO 274
LENGTH: 611
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..611
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..611
OTHER INFORMATION: Ceres Seq. ID 1009069
US-09-595-328C-274

Query Match 21.3%; Score 863; DB 6; Length 611;
Best Local Similarity 38.1%; Pred. No. 4.6e-61;
Matches 203; Conservative 95; Mismatches 183; Indels 52; Gaps 15;
QY 119 ASEQDSSEIMDANEOPQAKVTRTSIVFTGAEAPYAKSGGLGVDGCSLPALAAARGHRVMV 178
Db 65 SSRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLGDLGGLPPALAAARGHRVMTI 123
QY 179 MPRYLSSGSKNAYKALYTAHKKIPCFGGSHEVTFEYRDNDVDFVDPHPSY-----H 233
Db 124 CPYI-----DQYKDAWDTCVVQIKVGDKNVRFHFCHYKRGVDRVDFVDPHIFLAKVVG 177
QY 234 RPSG-LYGDNFGA-FGDNQRYTLLCYAAACEAPLILELGG--YI---YGQNCMFVNDWH 286
Db 178 KTSKIKYGPITGVYNDNQLRFSLLCQAALAPQVNLNLSKYSFGPYGEDVVFVANDWH 237
QY 287 ASLVPVLLAAKYRPGYVYRDSRSTLVHNLAHQGVPEASTYPDLGLPPEWYGALEWVFE 346
Db 238 TALLPCYLKSMYQSRGYVMNAKVFCIHNIAVQGRFAFDYDYLNLPISEKSSFDF---- 293
QY 347 WARRHALD-----KGEAVNFKAGVYADRTVTSOGYSWE-VTTAEGQGLNELSSR 399
Db 294 -----MDGYEKPVKGRKINMKAAILEAHRLVTVSPYQAELISGVDRGVELHLYL--R 345
QY 400 KSVLNGIVNGIDINDNPTDKCLPHHYSVDLLS-GKAKCKAELOKELGLPVREDVPLIG 458
Db 346 MKTVSGIINGMDVQWENPSTDKVIDIKYDITVTDAKPLIKEALQAAVGLPVRDVPVIG 405
QY 459 FIGRLDYOKGIDILKMAIPELMREDVQVMLGSDPIFEGWMRSTESSYKDKFRGWGFS 518
Db 406 FIGRLEQKSDILVEAISKFMGLNVQVILGTGKKKMEAQILELEKFPKAVGVAKFN 465
QY 519 VPVSHRTACDILLMPSRFEPCGLNQLYAMQYGTVPVHGTGGLRDTVET-----FNP 572
Db 466 VPLAHMITAGADFIVPSRFEPCGLIQLHAMRYGTVPVASTGGLVDTVKDGYTGFIHR 525
QY 573 FGAKGEGTGWASPLTVDKMLWALRTAMSTFREHKPSWEGLMKRGMTKDHTW 625
Db 526 FNVKCE-----VVDPPDVIATAKAVTRAVAVY--GTSAMQEMVKNCMQDQFSW 571

RESULT 14
US-09-595-328C-276
Sequence 276, Application US/09595328C
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-0946P
CURRENT APPLICATION NUMBER: US/09/595,328C
NUMBER OF SEQ ID NOS: 3107
SEQ ID NO 276
LENGTH: 528
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..528
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..528
OTHER INFORMATION: Ceres Seq. ID 1009071
US-09-595-328C-276

Query Match 21.3%; Score 862.5; DB 6; Length 528;
Best Local Similarity 38.9%; Pred. No. 4.1e-61;
Matches 199; Conservative 89; Mismatches 173; Indels 51; Gaps 14;

```
QY 140 STVFVTGEAARYAGSGGLGDDVCSLPIALAAARGHRVMMVMPRYLNGSSDKNYAKALYAK 199
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 2 SVIFIAEVPWSKGTGLGDLGGLPPALAAARGHRVMTICPRY-----DQYKDAMDTCV 55
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 200 HKIKFCGSGSHVTFHEYRDNDVDFVDPHSY-----HRPGS-lyGDNFGA-FGDNQFR 252
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 56 VVOIKVGDKVKNRFEHCYKRGVDVDFVDPHPIFLAKVVGKGTGSKIYGPITGVYDNDNQLR 115
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 253 YTLCAACEAPLILELGG--YI---YGONCMFVNDHSLVPLLAAKARYPYGVYRDS 307
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 116 FSLCQAALEAPQVNLNASSKFSFGPYGEDVVFVANDHWTALPCYLKSMYOSRGVYVNA 175
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 308 RSTLVHNLAHOGVPEASTYDPLGLPPEMYGALEWVFPWARHALD-----KGEAVNF 361
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 176 KVVFCIHNIAYGAFDDYSLLNLPISFKSSFDE-----MDGYEKPVKGRKINW 225
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 362 LKAVVYADRIYVTSQGSWE-VTTAEGGQGLNELLSRKS SVLNGIDINDNPTD 420
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 226 MKAATLEAHRVLTSPYQAQLISGVDGRGVELHKYL--RMKTVSGIINGMDVQEWNPSTD 283
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 421 KCLPHYSVDLDS-CKAKCAELQELGLPVREDYPLIGFGRLDYQKGIDILKNAIPEL 479
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 284 KYIDKYDITTVTDAKPLIKELQAAGLVPRDVPVIGFGRLEEQKGSIDLVEAISKF 343
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 480 MREDYQVFMVLSGDPFIFEGMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFE 539
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 344 MGLNVQVILGTGKKKMEAQILELEKPPGKAVGAKFNVPLAHMITAGADFIIVPSRFE 403
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 540 PGLNQLYAMQYCTVPVHGTGGLRDTVET-----FNPFGAKGEGTGWAFSPLTVDKM 593
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 404 PCGLIQLHAMRYGTVPVASTGGLVDTVKDGYTGFHIGRFNVKCE-----VVPDDVVIAT 458
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 594 LWALRTAMSTEREHKKPSNEGLMKRGMTKDHTW 625
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 459 AKAVTRAVAVY--GTSAMQEMVKNCMQDQFSW 488
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
```

```
RESULT 15
US-10-044-543-18
; Sequence 18, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthese Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches
; FILE REFERENCE: 1144D
; CURRENT APPLICATION NUMBER: US/10/044,543
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 614
; TYPE: PRP
; ORGANISM: Typha latifolia
US-10-044-543-18
```

```
Query Match 21.3%; Score 860; DB 7; Length 614;
Best Local Similarity 36.0%; Pred. NO. 8.1e-61;
Matches 201; Conservative 97; Mismatches 191; Indels 70; Gaps 14;

QY 108 AEDSDSIIVASEQDSEIMDAN-----EQPAKVT-----RSIVFTGEAAPYAKSGG 156
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 46 SQNSVDLQVTKAKTTSQKNGCGSTVNLPRPRAVICGKGMNLVFGAEMAPWSKGTG 105
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 157 LGDVCGSLPIALAAARGHRVMMVMPRYLNGSSDKNYAKALYAKIKPCFGGSHEVTFPH 216
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 106 LGDVLGGLPPALAAARGHRVMTAPRY-----DQYDAMDTDALVELKVGDRCEVTFPH 159
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 217 EYRDNDVDFVDPHSY-----HRPGSLYGDNFGA-FGDNQFRYTLACYAAACEAPLILEL 269
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
```

```
Db 160 CYKRGVDVDFVDPHMFARVCKTGKIYGPNTGYDQNLRFSLQAALEAPRIINL 219
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 270 GGY-----LYGONCMFVNDHSLVPLLAAKARYPYGVYRDSRSTLVTHNLAHOGVERPA 324
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 220 NNSDSFGPYGEDVIFICNDHMTSLPCYLKSMYHPRGIYKNAKVAFCIHNISYQGRFSP 279
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 325 STYDPLGLPPEMYGALEWVFPWARHALD-----KGEAVNFKGAVVYADRIYVTSQ 378
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 280 SDFEFLNLPENFKSSFSFI-----DGYNKPVKGMKINMKAGILESDFVFTVSPY 329
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 379 YSWEVTTA-EGGQGLNELLSRKS SVLNGIDINDNPTDCKLPHYSVDD-LSGKA 436
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 330 YAQELLSGEERGVELDNIL--RVTSITGIVNGMDVNEWNPLTDKYISVNYDAKPYMEAKP 387
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 437 KCKAELQELGLPVREDYPLIGFGRLDYQKGIDILKNAIPELMREDYQVFMVLSGDPITF 496
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 388 LNKEALQAESWLACRQGHPCNCIHGRLEEQKGSIDLAASIPEDINDENVQLIILGTGKKEM 447
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 497 EGMWRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFPCGCLNQLYAMQYGTVPV 556
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 448 ENQLESMEEMPDKYRAVMKFNAPLAHQMTAGADIIVIPSRFPCGLIQLQGMQYGTPSA 507
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 557 VHGTCGLRDTVETFPFGAKGEGTGWAFSPLT-----VDKMLWALRTAMSTFRE 606
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 508 CSSTGGLVDTVK-----EGKTGFHMGPPFSABCEVVDLSDVKKVVTTVKRALKVI-- 556
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 607 HKPSWEGLMKRGMTKDHTW 625
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 557 GTPAFEDMIKNCMAQDLSW 575
      |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
```

Search completed: March 28, 2002, 16:29:23
Job time: 577 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 15:01:15 ; Search time 50.54 Seconds
(without alignments)
336.614 Million cell updates/sec

Title: US-09-674-824-2
Perfect score: 4044
Sequence: 1 MAATGVGAGCLAPSVRLRAD.....SDGSLSVRTAETIRNQLVTL 756

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3612	89.3	671	4	US-09-196-390-2
2	2749	68.0	649	4	US-09-192-909-2
3	2557.5	63.2	583	3	US-08-941-445A-13
4	2390.5	59.1	539	3	US-08-941-445A-21
5	1992	49.3	641	4	US-08-836-567-10
6	1718.5	42.5	459	4	US-08-836-567-4
7	1215	30.0	767	4	US-08-836-567-8
8	1205.5	29.8	698	3	US-08-941-445A-11
9	1184	29.3	558	4	US-08-836-567-6
10	1148.5	28.4	799	4	US-09-196-390-6
11	1058	26.4	669	3	US-08-941-445A-9
12	877	21.7	609	3	US-08-941-445A-7
13	865	21.4	533	3	US-08-941-445A-5
14	678.5	16.8	477	1	US-07-735-065-2
15	678.5	16.8	477	1	US-08-469-202-12
16	678.5	16.8	477	2	US-08-484-434C-12
17	556	13.7	677	4	US-08-836-567-2
18	556	13.7	1197	4	US-08-836-567-12
19	556	13.7	1230	2	US-08-968-542C-35
20	551	13.6	735	4	US-09-115-704-2
21	492	12.2	1674	2	US-08-968-542C-12
22	176	4.4	64	2	US-08-470-720-15
23	158	3.9	31	2	US-08-572-951-26
24	146.5	3.6	79	2	US-08-470-720-13
25	138	3.4	33	2	US-08-572-951-5
26	133	3.3	31	2	US-08-572-951-25
27	133	3.3	59	2	US-08-470-720-14

28 118.5 2.9 2205 1 US-08-093-453B-2 Sequence 2, Appli
29 114.5 2.8 666 4 US-08-961-083-2 Sequence 2, Appli
30 110.5 2.7 587 1 US-08-398-008A-2 Sequence 2, Appli
31 110.5 2.7 587 2 US-08-893-333-2 Sequence 2, Appli
32 110.5 2.7 1045 2 US-08-553-436A-6 Sequence 6, Appli
33 109.5 2.7 682 3 US-08-481-435-6 Sequence 6, Appli
34 107.5 2.7 6095 4 US-09-144-085-2 Sequence 2, Appli
35 107 2.6 1054 1 US-08-356-354-4 Sequence 4, Appli
36 107 2.6 1054 2 US-08-778-656-4 Sequence 4, Appli
37 106.5 2.6 702 4 US-09-232-200-102 Sequence 102, App
38 106.5 2.6 702 4 US-09-232-197-102 Sequence 102, App
39 106 2.6 382 3 US-08-660-645A-7 Sequence 7, Appli
40 106 2.6 382 3 US-09-298-718-7 Sequence 7, Appli
41 106 2.6 382 4 US-09-546-969-7 Sequence 7, Appli
42 106 2.6 382 4 US-08-980-832-5 Sequence 5, Appli
43 103.5 2.6 495 1 US-07-841-997A-2 Sequence 2, Appli
44 103.5 2.6 495 1 US-08-290-301-2 Sequence 2, Appli
45 103.5 2.6 495 2 US-08-588-983-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-196-390-2
; Sequence 2, Application US/09196390
; Patent No. 63071125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Luticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., C/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 671 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-196-390-2

Query Match 89.3%; Score 3612; DB 4; Length 671;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TOPPLDAGVGLAPDLLLGGIAEDSDSIIVAASEODSEIMDANEQPOAKVTRISVFT 145
DB 1 TOPPLDAGVGLAPDLLLGGIAEDSDSIIVAASEODSEIMDANEQPOAKVTRISVFT 60
QY 146 GEAPYAKSGGLGDCVCSLPALAAHGRVVMVMPRYLNGSSDKNAKALYTAHKIKIPC 205
DB 61 GEAPYAKSGGLGDCVCSLPALAAHGRVVMVMPRYLNGSSDKNAKALYTAHKIKIPC 120
QY 206 FGSHEVTFPHEYRDNDVWVVDHPSYHRPGSLYGNFGAGNQRYTLLCYAAACEAPL 265
DB 121 FGSHEVTFPHEYRDNDVWVVDHPSYHRPGSLYGNFGAGNQRYTLLCYAAACEAPL 180
QY 266 ILEGGYIYGONCMFVVDHWSLVPVLLAAKVRPYGVYRDSRSTLVIHNLAHQGVEPAS 325
DB 181 ILEGGYIYGONCMFVVDHWSLVPVLLAAKVRPYGVYRDSRSTLVIHNLAHQGVEPAS 240
QY 326 TYPDLGLPPPEWYGALEWVFPPEWARRHALDKGEAVNFKGAVVTADRIVTYSQGSWEVTT 385
DB 241 TYPDLGLPPPEWYGALEWVFPPEWARRHALDKGEAVNFKGAVVTADRIVTYSQGSWEVTT 300
QY 386 AEGQGLNELLSRKSVLNGIVNGIDINWNPPTDKCLPHYSVDDLSGKAKKAELOKE 445
DB 301 AEGQGLNELLSRKSVLNGIVNGIDINWNPPTDKCLPHYSVDDLSGKAKKAELOKE 360
QY 446 LGIPVREDVPLIGFGRDLYQKGTDLTKMAIPELMREDOVFVMLGSDPIFEGWMRSTES 505
DB 361 LGIPVREDVPLIGFGRDLYQKGTDLTKMAIPELMREDOVFVMLGSDPIFEGWMRSTES 420
QY 506 SYKDKFRGWGVSFVPSVSHRITAGCDILLMPSRFPCGLNGLYAMQYGTVPVHVGTGGLRD 565
DB 421 SYKDKFRGWGVSFVPSVSHRITAGCDILLMPSRFPCGLNGLYAMQYGTVPVHVGTGGLRD 480
QY 566 TVTENPFGAKGEBGTGWAESPLTVDKMLWALRTAMSTFHEHKPSWEGLMKRGMTKDHTW 625
DB 481 TVTENPFGAKGEBGTGWAESPLTVDKMLWALRTAMSTFHEHKPSWEGLMKRGMTKDHTW 540
QY 626 DHAPSSRSRSGSPSWNTNPTSCRRGLGRSKCESPSALKTSRSSFRGPEGYPCTLRCPATV 685
DB 541 DHAPSSRSRSGSPSWNTNPTSCRRGLGRSKCESPSALKTSRSSFRGPEGYPCTLRCPATV 600
QY 686 ESQCACLLWFAGSRTYDGCAGAAAVTASGGRLQFQWIRKGAAGLWLTAKHHSDGSLSVRV 745
DB 601 ESQCACLLWFAGSRTYDGCAGAAAVTASGGRLQFQWIRKGAAGLWLTAKHHSDGSLSVRV 660
QY 746 TAEIRNQLVTL 756
DB 661 TAEIRNQLVTL 671

RESULT 2
US-09-192-909-2
; Sequence 2, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Froberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; TITLE OF INVENTION: starch synthases from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York

; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02527
; FILING DATE: 16-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-192-909-2

Query Match 68.0%; Score 2749; DB 4; Length 649;
Best Local Similarity 82.5%; Pred. No. 1-7e-247;
Matches 518; Conservative 34; Mismatches 62; Indels 14; Gaps 4;
QY 4 TGVGAGCLAPSRLRADPATAARASACVWRA---RLRRLARGRYVAELSRGPAARPAQ 60
DB 14 SAVGAACLL-----LARAAMPAAVGDRAAPRRRLQRLRRRCVAELSRGPAARPMPP 65
QY 61 QQLAPLVPGLFAPPAPAPASPTOPPLPDAGVAGELAPDLLLGLIAEDSDSIIVAAAS 120
DB 66 ALLAPLVPGLFLA-PPAEPTEGEPALTPPPVDPDAGLVGLVE--PEGIAESIDNTVVVAS 122
QY 121 EQDSEIMDANEQPOAKVTRISIVFTGEAPYAKSGGLGDCVCSLPALAAHGRVVMVMP 180
DB 123 EQDSEIVVGKEQARAKVTQNVFVTGEASPYAKSGGLGDCVCSLPALAAHGRVVMVMP 182
QY 181 RYLANGSDKNYAKALYTAHKIKIPCTFGGSHSEVTFHEYRDNDVWVVDHPSYHRPGSLY 240
DB 183 RYLANGSDKNYANAFYTEKHIRIPCEGGEHEVTFHEYRDSVDVWFYDHPHYHRPGLNYG 242
QY 241 DNFGAFGDNQFRYTLCLCYAAACEAPLILELGGYIYGONCMFVVDHWSLVPVLLAAKVRP 300
DB 243 DKFGAFGDNQFRYTLCLCYAAACEAPLVLELGGYIYGONCMFVVDHWSLVPVLLAAKVRP 302
QY 301 YGYRDSRSTLVIHNLAHQGVEPASTYDGLGPPPEWYGALEWVFPPEWARRHALDKGEAVN 360
DB 303 YGYVKDSRSLSLVIHNLAHQGVEPASTYDGLGPPPEWYGALEWVFPPEWARRHALDKGEAVN 362
QY 361 FLGAVVTADRIVTYSQGSWEVTTAEGGGLNELLSRKSVLNGIVNGIDINWNPPTD 420
DB 363 FLGAVVTADRIVTYSQGSWEVTTAEGGGLNELLSRKSVLNGIVNGIDINWNPPTD 422
QY 421 KCLPHYSVDDLSGKAKKAELOKELGCPVREDVPLIGFGRDLYQKGTDLTKMAIPELM 480
DB 423 KCIPCHYSVDDLSGKAKKAELOKELGCPVREDVPLIGFGRDLYQKGTDLTKMAIPELM 482
QY 481 REDVQFVNLGSGDPIFEGWMRSTESSYKDKFRGWGVSFVPSVSHRITAGCDILLMPSPFEP 540
DB 483 REDVQFVNLGSGDPIFEGWMRSTESSYKDKFRGWGVSFVPSVSHRITAGCDILLMPSPFEP 542

QY 541 CGLNOLYAMOGTVPVVHCTGGGLRDTVETNPFGAKGEGTGWAFSPLTVDKMLALRTA 600
DB 543 CGLNOLYAMOGTVPVVHCTGGGLRDTVETNPFGAKGEGTGWAFSPLTVDKMLALRTA 602
QY 601 MSTFREHKPSWEGMLKRGMTKDHTWDHA 628
DB 603 ISTYREHKPSWEGMLKRGMSKDFTDHA 630

RESULT 3
US-08-941-445A-13
; Sequence 13, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-941-445A-13

Query Match 63.2%; Score 2557.5; DB 3; Length 583;
Best Local Similarity 86.7%; Pred. No. 1e-229;
Matches 476; Conservative 25; Mismatches 45; Indels 3; Gaps 2;

QY 45 VAEISREGPAAPAOQOOLAPPLVPGFLAPPAPPAQSPAPTPPLPDAGVGLAPDLLL 104
DB 2 VAEISREGPAAPRLPALLAPPLVPGFLA-PPAETGEPASTPPVPDAGLDGLG--LEP 58
QY 105 EGIADSDISIIVASEQDSEIMANEQOQAKVTRISIVFTGEAAPIYAKSGGLGVDVCSL 164
DB 59 EGIADSDINTVIVASEQDSEIIVGKEQARAKVTQSIVFTGEASPIYAKSGGLGVDVCSL 118
QY 165 PIALAARGHRVMVMPRYLNGSDSKNYAKALYKIKIPCGGSHVETFFHEYRDNDVW 224
DB 119 PVALAARGHRVMVMPRYLNGSDSKNYANAFYTEKHIRIPCGGHEVETFFHEYRSDVDW 178
QY 225 VFVDHPSYHRPGSLYGDNGFAGDNOFRYTLCLYAAACEAPLILELGGYIYQNCMFVND 284
DB 179 VFVDHPSYHRPGSLYGDNGFAGDNOFRYTLCLYAAACEAPLILELGGYIYQNCMFVND 238

QY 285 WHASLVPVLLAAKRYPGVYRDSRSLTVIHNLHQGVPEASTYPTDLGLPPEWYGALEWVF 344
DB 239 WHASLVPVLLAAKRYPGVYKDSRSLTVIHNLHQGVPEASTYPTDLGLPPEWYGALEWVF 298
QY 345 PEWARHDLKGEAVNPLKGAVTADRIYTVSOGYSWEVTTAEGGOGNELLSSRSKSVLN 404
DB 299 PEWARHDLKGEAVNPLKGAVTADRIYTVSOGYSWEVTTAEGGOGNELLSSRSKSVLN 358
QY 405 GIVNGIDINDNPTDKCLPHHYSVDDLSGKAKCAELQKELGLPVREDVPLIGFIGRLD 464
DB 359 GIVNGIDINDNPTDKCIPCHYSVDDLSGKAKCAELQKELGLPIRPDVPPLIGFIGRLD 418
QY 465 YQKIDILIKMAPELMDREDVQFVNLGSGDPIFEGNMRSTESSYKDKFRGWGFSVPVSHR 524
DB 419 YQKIDILIQIITPDLMDREDVQFVNLGSGDPELEDNMRSTESIFKDKFRGWGFSVPVSHR 478
QY 525 ITAGCDILLMPSEFPCGLNOLYAMOGTVPVVHCTGGGLRDTVETNPFGAKGEGTGW 584
DB 479 ITAGCDILLMPSEFPCGLNOLYAMOGTVPVVHCTGGGLRDTVETNPFGAKGEGTGW 538
QY 585 FSPLTVDKM 593
DB 539 FAPLTENM 547

RESULT 4
US-08-941-445A-21
; Sequence 21, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-941-445A-21

Query Match 59.1%; Score 2390.5; DB 3; Length 539;
Best Local Similarity 88.4%; Pred. No. 3.3e-214;
Matches 444; Conservative 23; Mismatches 32; Indels 3; Gaps 2;

```

QY 95 VGELA-PDLLL--EGTAEDSIDSIIVAASQDSEINDANEPQOAKVTRISIVFTVGTGAAPY 151
Db 2 VAELSREDLGEPIGAEGSIDNTVVASBQDSEIVVKGQAKVQTQSVFTVGTGAAPY 61
QY 152 AKSGGLDVCGLPILAAARGHRVVMVMPRYLNGSDKNYAKALYAKHIIKIPCFGSGHE 211
Db 62 AKSGGLDVCGLPILAAARGHRVVMVMPRYLNGSDKNYAKALYAKHIIKIPCFGSGHE 121
QY 212 VTFEHEYRDNDVWFVVDHPHSYHRPGSLYGNFAGFAGDNQPRYTLILCYAACEAPLILELGG 271
Db 122 VTFEHEYRDNDVWFVVDHPHSYHRPGSLYGNFAGFAGDNQPRYTLILCYAACEAPLILELGG 181
QY 272 YIYGQNCMFVYNDWHSALVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYVDLG 331
Db 182 YIYGQNCMFVYNDWHSALVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYVDLG 241
QY 332 LPPEWYGALEWVPEWARRHALDKGEAVNFKGAVTADRIVTYSQYSGVSEVTTAEGGQ 391
Db 242 LPPEWYGALEWVPEWARRHALDKGEAVNFKGAVTADRIVTYSQYSGVSEVTTAEGGQ 301
QY 392 LNELLSRKSVLNGVINGIDINDWNPTTKCLPHHSYVDDLSGKAKCKAELQKELGLPVR 451
Db 302 LNELLSRKSVLNGVINGIDINDWNPTTKCLPHHSYVDDLSGKAKCKAELQKELGLPVR 361
QY 452 EDVPLIGFGRDYQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKF 511
Db 362 PDVPLIGFGRDYQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKF 421
QY 512 RGNWGFSPVSHRITAGCDIILLMPSRPEPCGLNOLYAMQYGVVHGTGGLRDTVETEN 571
Db 422 RGNWGFSPVSHRITAGCDIILLMPSRPEPCGLNOLYAMQYGVVHGTGGLRDTVETEN 481
QY 572 PFGAKGEGTGWAFSPVLTVDKM 593
Db 482 PFGAKGEGTGWAFSPVLTVDKM 503

RESULT 5
US-08-836-567-10
; Sequence 10, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens.
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836, 567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids.
; TYPE: amino acid
; TOPOLOGY: linear.
; MOLECULE TYPE: protein
; US-08-836-567-10

Query Match 49.3%; Score 1992; DB 4; Length 641;
Best Local Similarity 67.5%; Pred. No. 6.1e-177;
Matches 360; Conservative 80; Mismatches 87; Indels 6; Gaps 4;

QY 102 LLEGTAEDS--IDSIIVAASEQD--SEIMDANEQOAKVTRISIVFTVGTGAAPYKSGGL 157
Db 90 LIPHSVAGDATWVESHDIIVANDRDLDSEDETEEMEETPIKLTFTNIIIVTAEAPYSKTGGL 149
QY 158 GDVCGSLPTALAAARGHRVVMVMPRYLNGS--SDKNYAKALYAKHIIKIPCFGSGSHEVTFPH 216
Db 150 GDVCGSLPTALAAARGHRVVMVMPRYLNGSPSDEKYNANAVDLDVRAIVHCFGDAQEAFAFH 209
QY 217 EYRDNDVWFVVDHPHSYHRPGSLYGNFAGFAGDNQPRYTLILCYAACEAPLILELGGYIYQ 276
Db 210 EYRAGVDVWFVVDHPHSYHRPGSLYGNFAGFAGDNQPRYTLILCYAACEAPLILELGGYIYQ 269
QY 277 NCMFVNDWHSALVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYVDLGGLPPEW 336
Db 270 KCLFLANDWHAALVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYVDLGGLPPEW 329
QY 337 YGALEWVPEWARRHALDKGEAVNFKGAVTADRIVTYSQYSGVSEVTTAEGGQGLNEL 396
Db 330 YGAVWEIFPTWARAHALDTGETVNLKGAIVADRIVTYSQYSGVSEVTTAEGGQGLNEL 389
QY 397 SSRKSVLNGVINGIDINDWNPTTKCLPHHSYVDDLSGKAKCKAELQKELGLPVRDVL 456
Db 390 SSRKSVLNGVINGIDINDWNPTTKCLPHHSYVDDLSGKAKCKAELQKELGLPVRDVL 449
QY 457 IGFGRDLRYQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKFERGWG 516
Db 450 IGFGRDLRYQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKFERGWG 509
QY 517 FSPVSHRITAGCDIILLMPSRPEPCGLNOLYAMQYGVVHGTGGLRDTVETENPFGAK 576
Db 510 FNPVSHRITAGCDIILLMPSRPEPCGLNOLYAMRYGTIPVHSTGGLRDTVETENPFGAK 569
QY 577 G-BEGTGWAFSPVLTVDKMALRTAMSTFPEHKKPSWEGMLMKRGTWKDHTWDHA 628
Db 570 GIGEGTGWAFSPVLTVDKMALRTAMSTFPEHKKPSWEGMLMKRGTWKDHTWDHA 622

RESULT 6
US-08-836-567-4
; Sequence 4, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York

```

```

; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836.567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-836-567-4

```

```

Query Match 42.5%; Score 1718.5; DB 4; Length 459;
Best Local Similarity 70.6%; Pred. No. 1.1e-151;
Matches 303; Conservative 62; Mismatches 61; Indels 3; Gaps 2;

QY 203 IPCFGGSHVETFEHYRDNVDVFDHPSYHRSGLYDNGFAGDQNFRTLLCYAAACE 262
DB 14 VHCFGDAQVAFVETRYAGVDVDFVHSSYRRTGYDYGAFGDNQFRFTLLSHAACE 73
QY 263 APLILEGGYIYQNCMFVNDVHSLVPLLAARYPYRGRDSTRSLVHNLAHQGYE 322
DB 74 APLVPLGGTYGKCLFLANDCNALVPLLAARYPYRGRDSTRSLVHNLAHQGYE 133
QY 323 PASTYDPLGLPPWYGALEWFPWARRHALDKGEAVNFKGAVTVADRVTVSQYSWE 382
DB 134 PAVTYNNLGLPPWYGALEWFPWARRHALDKGETVNVLKGAIADVADRVTVSQYSWE 193
QY 383 VTTAEGGQGLNELLSRKSVLNGVINGIDINDWNPPTDKCLPHHYSVDLSS--GRAKCKA 440
DB 194 ITTPEGGYGLHELLSRQSVLNGVINGIDINDWNPSTDEHIAHSYINDLSPGKVQCKT 253
QY 441 ELQKELGLPREDVPLIGFGRDLYOKGIDLIKMAIPELMREDVQFVMLGSGDPIFGWM 500
DB 254 DQKELGLPIRDCPLIGFGRDLYOKGIDLIKMAIPELMREDVQFVMLGSGDPIFGWM 313
QY 501 RSTESYKDKFGWGFSPVSHRTAGCDILLMPSPRCPGLNOLYAMQYGTVPVHVHT 560
DB 314 RHTENLFKDKFRWGVFNVPVSHRTAGCDILLMPSPRCPGLNOLYAMRYGTIPVHT 373
QY 561 GGLRDTVTNFPFGAKGE--EGTGWAFSPLTVDKMLWALRTAMSTPREHKPSWEGMLKRG 619
DB 374 GGLRDTVKDFNPYAQEGGEGTGWTFSPLTSEKLFDTLKLARTYTEHKSSWEGMLKRG 619
QY 620 TKDHTWDHA 628
DB 434 GRDYSWENA 442

```

```

RESULT 7
US-08-836-567-8
; Sequence 8, Application US/08836567

```

```

; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836.567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-836-567-8

```

```

Query Match 30.0%; Score 1215; DB 4; Length 767;
Best Local Similarity 45.7%; Pred. No. 2.1e-104;
Matches 264; Conservative 78; Mismatches 188; Indels 48; Gaps 12;

QY 79 PAQSPAPTQPL-----PDAGVGLAPDLLEGTA-----EDSIDSIIVAA 120
DB 195 PPTPKSQEFTLLDVNSRKSLVDYVPGKKIQSYMPSLRKSSASHVEQNEENLEGSAAE 254
QY 121 BQDSEIDANBQPAKV---TRSIIVTGAAPYAKSGGLGDCVCSLPALAAARGHRVM 177
DB 255 EETEDPVNIDEKPPPLAGTNVMIIVASCAPWSKTGGGLGDVAGALPKALARRHRVM 314
QY 178 VMPRYLNGSDKNYAKALYTAHKIKPCFGSGSEVTEFFHEYRDNVDVFDV--HPSYHRPG 236
DB 315 VAPRYDNPPEPQDSG---VRKIYKVD--QGDVEVTYFQAFIDGVDVFDIDSHMFRHGN 368
QY 237 SLYGDNGFAGDQNFRTLLCYAACEAPLILEGGYIYQ--NCMFVNDVHSLVPLVLLA 295
DB 369 NIYGN---RVDILKRWLFCKAIEVPHVPCGVCYGDGNLVFTIANDWHTALLPVYK 425
QY 296 AKYRPGYVYRDSRSLVIHNLAHQGVPEASTYDPLGLPPEWYGALEWFPWARRHALDK 355
DB 426 AYVRDNGIMNYTRSVLVHNLAHQGVPEASTYDPLGLPPEWYGALEWFPWARRHALDK 355
QY 356 GEAVNFKGAVTVADRVTVSQYSWEVTTAEGGQGLNELLSRKSVLNGVINGIDINDW 415
DB 426 AYVRDNGIMNYTRSVLVHNLAHQGVPEASTYDPLGLPPEWYGALEWFPWARRHALDK 355

```

Db 478 GEHNFIFAAGLKTADRVVTVSHGYSWELKTSQGGMLHGOIINENDKWLOGIVNGIDTKEW 537
Qy 416 NPPTDKCLPH---HYSVDL-SGKAKCAELQKELGLPVRDVPVLPVIGRDLQKQID 470
Db 538 NPELDVHLQSDGVMYNSLDTLQTKPKQCKAALQKELGLPVRDVPVLPVIGRDLQKQVD 597
Qy 471 LIKMAPELMREDVQVFMVLMGSDPIFEGWMRSTESSYKDKFRGVGVFVPSVSHRITAGCD 530
Db 598 LIAEASAMWQDVQLVLTGTRDLEQMLRQFECHNDKIRGVGVFVPSVSHRITAGAD 657
Qy 531 ILLMPSRFEPCGLNOLYAMQYCTVPVHGTGGLRDRVTFNPEGAKGEGTGWAFSPLTV 590
Db 658 ILLMPSRFEPCGLNOLYAMQYCTVPVHGTGGLRDRVTFNPEGAKGEGTGWAFSPLTV 590
Qy 591 DKMLWALRTAMSTFRHFKPSWGLMKRGMTKDHTWDHA 628
Db 715 SOLIHALGNCLLYREYKKSMEGIGTQRCMTQDLSWDA 752
RESULT 8
US-08-941-445A-11
: Sequence 11, Application US/08941445A
: Patent No. 6107060
: GENERAL INFORMATION:
: APPLICANT: Keeling, Peter
: APPLICANT: Guan, Hanping
: TITLE OF INVENTION: Starch Encapsulation
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Circle
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08941.445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 698 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-941-445A-11
Query Match 29.8%; Score 1205.5; DB 3; Length 698;
Best Local Similarity 43.0%; Pred. No. 1.4e-103;
Matches 283; Conservative 83; Mismatches 221; Indels 71; Gaps 17;
Qy 12 APSVRLRADPA-----TAASACVVRRLRLRGRVYAEVLSREGPAARPAQOQOOLAPP 66
Db 56 AASVRAAAAPAGSEAEAKSSS-SQAGAVQGSTAKAVDSASPPNPLTSPAPKQSQSAM 114
Qy 67 L--VPGFLAPPAPPAQSPAPTOPPLP-----DAG-----VGLAPDLLLEGIAEDSI 112

Db 115 QNCTSGSSASTAAPVSGPKADHPSAPVTKREIDASAVKPEPAGDDARPVESIGIAE-PV 173
Qy 113 DSIIVAASEQDS-----EIDANEQQA-KVTRSIYFVTGEAAPYAKSGGLGDDVCGSLP 165
Db 174 DAKADAAPATDAASAPYDREDNEPGLAGNVMNVMNVVVAASECAPCKTGGGLGVGALP 233
Qy 166 IALAARGHRVWVMPRYLNGSSDKYAKALYAKHIKIPCFGSGSHEVYTFTHYRDNDVWY 225
Db 234 KALARGHRVWVTPRY-----GEYAEARDLGVRRYKAVAGQDSVTFYHSHYDGVDFV 287
Qy 226 FVDHPSY-HRPGSLYGDNFGAFGNQFRYTLCLCYAAEAPLILELGGYIYCO-NCMEVWN 283
Db 288 FVEAPFERHNNIYG---GERLDLKRMIILFCKAAVEVPWYAPCGGTGVYDGGDLVFIA 344
Qy 284 DHASLVPLVLAARYPYGVYDRSRLVLIHNLHQGVPEASTYPDGLGPPWYGALEWV 343
Db 345 DMHTALLPVYKAYYRDNLGMOYARSVLIHNLHQGVPEASTYPDGLGPPWYGALEWV 343
Qy 344 FPEWARRHALDK-----GEAVNFKGAVVTADRVTVTSQGYSWETTTAGGGGLNEL 395
Db 395 -----EHYIDHFKLYDNIGGDHNSVFAAGLKTADRVTVTSVNGYMWELKTSEGGMLHDI 448
Qy 396 LSSRKSLVINGIDINDWNPPTDKCLPH---HYSVDL-SGKAKCAELQKELGLP 450
Db 449-INONDKWLOGIVNGIDSEWNPVAVDVHLSDDYTYTFETLDTGKQCKAALQRLGLQV 508
Qy 451 REDVPLIGFIRLDYQKIDILKMAIPELMRDVOFVLMGSDPIFEGWMRSTESSYKDK 510
Db 509 RDDVPLIGFIRLDYQKIDILKMAIPELMRDVOFVLMGSDPIFEGWMRSTESSYKDK 568
Qy 511 FRGVGFSVPVSHRITAGCDIILLMPSRFEPCGLNOLYAMQYCTVPVHGTGGLRDTVTF 570
Db 569 VRAVGFVSPLAHRITAGADILLMPSRFEPCGLNOLYAMQYCTVPVHGTGGLRDTVTF 628
Qy 571 NPEGAKGEGTGWAFSPLTVDKMLWALRTAMSTFRHFKPSWGLMKRGMTKDHTWDHA 628
Db 629 DPF---NDTGLGTFDRAENRMIDALSCLTTYRNYKESWRACRARGMAEDLSMDHA 683
RESULT 9
US-08-836-567-6
: Sequence 6, Application US/08836567
: Patent No. 6130367
: GENERAL INFORMATION:
: APPLICANT: Kossmann, Jens
: APPLICANT: Springer, Franziska
: APPLICANT: Abel, Gernot
: TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
: TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
: TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FISH & NEAVE
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/836,567
: FILING DATE: 24-JUL-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP95/04415
: FILING DATE: 09-NOV-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE P 44 41 408.0
: FILING DATE: 10-NOV-1994


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 609 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-941-445A-7

```

Query Match 21.7%; Score 877; DB 3; Length 609;
Best Local Similarity 38.5%; Pred. No. 4.7e-73;
Matches 205; Conservative 89; Mismatches 173; Indels 66; Gaps 15;

	QY	140	SIYFVTEAAPYAKSGGLGDVCSPALAAARGHRVVWVPYLNGSSDKNTAKALYTAK	199
	Db	84	NVYFGAEMAPWSKTGGLDVGGLGPMAANGHRVMYSIPRY-----DOYKDAWDTSV	137
	QY	200	HIKIPCGGSHEVTFTHHEYRDNDVDWFVDHPSY-----HPRG-SLYXGDNFGA-FGDNQFR	252
	Db	138	VAEIKVADRYERVRFHCYKRGVDRVFIDHPSPLEKVMWKTGEKIYGPDTGVYKDQMR	197
	QY	253	YTLICAAACEAPLILLEGTY-----IYGONCMFVNDWHASLPVLLAKARYPGYVRDS	307
	Db	198	FSLUCQAALAEAPRILLNNPNPYFKGYGDEVVFCVNDHTGTPLASYLKNNYPNGITYRNA	257
	QY	308	RSTLVTHNLAHQCVESPASTYPDGLPPEWYGALEWYFEWARRHALLD---KCEAVNFLKG	364
	Db	258	KVAFCITHNISYQGFAPEDYPELNLSERFSSEDFI-----DGYPDTPVEGRKINWMKA	310
	QY	365	AVVTADRIVTVSOGYSWE-VTTABGGQGGLNELLSRRKSVLINGIVINDINDNNPTDKCL	423
	Db	311	GILEADVLTVSPIYYAEELISGTARCELNDTN--RLTGITGITVGMWSEMDPSDKXI	368
	QY	424	PHHY-SVDDLSGRKAKCAELORELGPUVEDVPLIGFTGRLDYQKGIDILIKNAIPELMRE	482
	Db	369	TAKYDATTAIEAKALKNEALQAEAGLPVDRKIPLIAFIGRLBEOGPDFVMAAIIPELMQE	428
	QY	483	DVOFMLGSDPIIFEGWMKSTESSYKDKPRGWGVFSVPVSHRTTAGCDILLMPSRPEPCG	542
	Db	429	DVOIVLLGTYGKKFEKLKSMEEKYPGKYRAVVVKFNAPLAHLIMAGADVLAVPSRPEPCG	488
	QY	543	LNOLYAQMCTPVVPHVHGTCGLTRDTVETENPFNGAKBEEGTGWAFSPLTVD-----K	592
	Db	489	LIQLOMGXYGTPCACASTGGLVDTVI-----EGKTGFHWGRLSUVDCKVVEPSPDVKK	539
	QY	593	MLWALTAMSTFREHHKPSWEGMLMKRGMTKDHTDWHDAPSSTRSSSGP---SWTN	643
	Db	540	VAATLRKAIKVW-GTTPAYEEMVRNCMDQLW-----KGPANKWEN	579

```

RESULT 13
US-08-941-445A-5
; Sequence 5, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encap
; NUMBER OF SEQUENCES: 37

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,445A
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,855
FILING DATE: 30-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P.
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-941-445A-5

Query Match 21.4%; Score 865; DB 3; Length 533;
Best Local Similarity 38.0%; Pred. No. 5e-72;
Matches 205; Conservative 90; Mismatches 177; Indels 68; Gaps 16;

Qy	135	AKVTRTSIVFTVGEAAPYAKSGGLDVCGLSPIALAARGHRVWVMPRYLNGSSDKNYAKA	194
I			
Db	1	ASAGMNVVYGAEMAPMWSKTGGGLDVLGGLPAPMAANGHRVWVSPRY	54
I			
Qy	195	LYTAKHIIKIFCGSGSHBEVTFHEVRDNVDWVFVDHPSY	247
I			
Db	55	WDRSVNSEIKMGDGYETVRFFHCYKRGVDRFVDHPLFLERWCKTEEKLYGVPAGTDYR	114
I			
Qy	248	DNQFRYTLLYCAACEAPLILELGG--YI---YQNCNVFVYNDWHASLVPVLLAAKYPXG	302
I			
Db	115	DNQLRFSLLCQAALAPRILSLNNPNYPSPGYEDVVFVYCNWDHWTGPLSCYLKSNYQSHG	174
I			
Qy	303	VYRDSRSTLVIHNLAHOGVEPASTYPDLGLPPWYGALEWHVFEWARRHALDK--GEAV	359
I			
Db	175	IYRDAKTAFCINIIISYOGRAFSDYPELNIUPERFKSFDPI-----DGYEKPVEGRKI	227
I			
Qy	360	NFLKGAVTADRVITYVSQGSWE-VPTAEGQGGLNELLSRKSVLNGVNGIDINDWNP	418
I			
Db	228	NMKAGILEADRVLTYSPYAEELISGIARGCELDNIM--RLTGITGVNGMDVSEWDS	285
I			
Qy	419	TDKCLPHIYSVDD--LSKAKCKAEAKELGELGVREDVPLIGFTGRDLYQKGIDLIKMAIP	477
I			
Db	286	RDXYIAVKYDVSTAVEAKALNKEALQAEVLGPVDRNIPLVAFITGRLEEQKGPDMVMAAIP	345
I			
Qy	478	ELMR--EDVOFVMLGSDPIFEGNMRSSTESSYDKKRGWGVGFSVPVSHRITACGCDILLMP	535
I			
Db	346	QLMEMVEDQIVILGTGKKKFERMLSAEEKFPKGVRVYKFNAAAHIMAGAOLVAT	405
I			
Qy	536	SREPEPCNLQYAMQYGTVPVWHGTGGLRTVETFNFPFKAKEGEGTGAFAFSPLTVD----	591
I			
Db	406	SREPECGLILOQMRYGTPCCACASTGGLVDITII-----EGKTGFHWGRLSVDVCNV	456
I			
Qy	592	-----KMLWALRTAMSTFREHKPSWEGLMKMGKMTKDHTWDHAPSSTRSSSGP--SWTN	643
I			

Db 457 EPADVKKVATTLOAIAKV--GTPAYEEMVRNMIQDLSW-----KCPAKNWN 503

RESULT 14

US-07-735-065-2

Sequence 2, Application US/07735065

Patent No. 5349123

GENERAL INFORMATION:

APPLICANT: Christine K. Shewmaker

APPLICANT: David M. Stalker

TITLE OF INVENTION: Glycogen Biosynthetic Enzymes

TITLE OF INVENTION: In Plants

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Calgene, Inc.

STREET: 1920 Fifth Street

CITY: Davis

STATE: CA

COUNTRY: USA

ZIP: 95616

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 6.0.7

SOFTWARE: Microsoft Word 4.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/735,065

FILING DATE: 19910724

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/632,383

FILING DATE: 21-DEC-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: unassigned

FILING DATE: 16-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Lassen

REGISTRATION NUMBER: 31,845

NAME: Donna E. Scherer

REGISTRATION NUMBER: 34,719

REFERENCE/DOCKET NUMBER: CGNE 77-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (916) 753-6313

TELEFAX: (916) 753-1510

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-735-065-2

Query Match 16.8%; Score 678.5; DB 1; Length 477;

Best Local Similarity 34.1%; Pred. No. 1e-54;

Matches 170; Conservative 78; Mismatches 210; Indels 41; Gaps 12;

QY 141 IVFVTGAAPYAKSGGLGDCVCSLPTALAAARGHRVWVMPRY---LNGSSDKNYAKALYT 197

Db 3 VLVHCSEMFLLTKGLADVIGALPAAQIADGVDAARVLLPAFPDIRGVTDQVVSRRDT 62

QY 198 AKHIKIPCGSGSEVTFHFHYRNDVWFVDHPS-YHRPGSLYGD-NFGAFGDNQFRTL 255

Db 63 -----FAG--HITLLFGHYNGVGLYDAPHLVDYRPGSPYHDTNLFAYTNVLRFL 112

QY 256 LCVAACEAPLILEGGVIYQNCMFVNDWHSALVPVLLAAKYRPGYGVYRDSRSTLIHN 315

Db 113 LGWVGAE-----MASGLDPFRPVDVHHDHAGLAPAYLAARGP-----AKSVFTGHN 162

QY 316 LAHOGVEPASTYFDLGLPPEWYGALEKVFPEARRHALDKGEAVNFKLGAVVTADRTVTV 375

Db 163 LAYQGMFYAHMNDIQLP-----WSF-----FNIHGLEFNGQISFLKAGLYVADHITAV 211

QY 376 SQGISHEVTTAEGCGQLNELLSS--RKSVLNGIVNGIDINDNPTTDKCLPHHYSVDDLS 433

Db 212 SPTYAREITEPQFAYGMEGLLQQRHREGRLSGVLNGVDEKINSPEITDLLLLASRYTRDTLE 271

QY 434 GKACKAEQLQKELGVPREDVPLIGFTIGRLDYQKGLDLIKMAIPELMREDVQFVMLGSGD 493

Db 272 DKAENKRLQIANGLVKDDKVPFAVVRSLTSQGLDLVLEALPGLLEGGQLALLGAGD 331

QY 494 PIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFPFCGLNQLYAMQYGT 553

Db 332 PVLQEGFLAAAEYPGQVGVOIGYHEAFSHRIMGADVILVPSRFPFCGLTQLYGLKYGT 391

QY 554 VPVHGTGGLRDTYETFNFGKGEGBGTGWAFPLVDIKMLWALRTAMSTFRHKKPS-WE 612

Db 392 LPLVRRRTGLADTVSDCSLENLADGVASGFVEDSNWLLRAIRRAFVLW--SRPSLMR 449

QY 613 GLMKRGMTKDHTWDHAPSS 631

Db 450 FVORQAMDFSWQVAKS 468

RESULT 15

US-08-469-202-12

Sequence 12, Application US/08469202

Patent No. 5750875

GENERAL INFORMATION:

APPLICANT: STALKER, DAVID

TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: CALGENE, INC.

STREET: 1920 FIFTH STREET

CITY: DAVIS

STATE: CA

COUNTRY: USA

ZIP: 95616

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.1

SOFTWARE: Microsoft Word 5.1 (a)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,202

FILING DATE: 6-JUNE-95

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/016,881

FILING DATE: 11_FEB_1993

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Lassen

REGISTRATION NUMBER: 31,845

NAME: Donna E. Scherer

REGISTRATION NUMBER: 34,719

NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924

REFERENCE/DOCKET NUMBER: CGNE 93-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 916-753-6313

TELEFAX: 916-753-1510

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-08-469-202-12

Query Match 16.8%; Score 678.5; DB 1; Length 477;

Best Local Similarity 34.1%; Pred. No. 1e-54;

Matches 170; Conservative 78; Mismatches 210; Indels 41; Gaps 12;

Search completed: March 28, 2002, 16:21:44
Job time: 4829 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: March 28, 2002, 12:08:26 ; Search time 2297.02 Seconds
(without alignments)
12963.123 Million cell updates/sec

Title: US-09-674-824-1
Perfect score: 2771
Sequence: 1 cgcactccactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	601	21.7	988	11	BG343782
2	564	20.4	706	10	AV832913
3	494	17.8	988	10	BE418354
4	403	14.6	557	11	BG607242
5	393	14.2	562	10	AW678087
6	388	14.0	454	10	AV832689
7	380	13.7	408	11	BE482660
8	325	11.8	596	11	BE113316
9	320	11.6	588	10	BE434991
10	294	10.6	396	10	AU075475
11	276	10.0	401	11	BG050773
12	250	9.0	564	10	AW932199

13	225.2	8.1	618	10	AW759569
14	223.4	8.1	623	10	AW759946
15	216.4	7.8	335	11	C19776
16	215.2	7.8	411	10	AW438020
17	205	7.4	217	11	BF478674
18	195.4	7.1	592	10	AW599325
19	194.8	7.0	294	10	AW418330
20	182.2	6.6	243	10	AW749940
21	168.6	6.1	548	11	BF424294
22	165.2	6.0	601	11	BG453184
23	159	5.7	632	10	AW678169
24	156	5.6	729	11	BG599615
25	154.8	5.6	773	11	BG511175
26	154.4	5.6	696	11	BG351920
27	148.8	5.4	734	11	BG350240
28	145.8	5.3	625	10	BE347900
29	145.4	5.2	649	11	BE406798
30	144.2	5.2	455	10	BE555965
31	141.4	5.1	143	10	BE425099
32	141.2	5.1	753	11	BF460244
33	140	5.1	589	10	AW569428
34	136.8	4.9	614	10	AV524142
35	136.2	4.9	446	11	BG647513
36	131	4.7	568	11	BG350144
37	130	4.7	679	11	BE406287
38	128.6	4.6	522	11	BG652154
39	127.4	4.6	754	11	BG351232
40	126	4.5	565	11	BG154708
41	125.2	4.5	635	10	AV441053
42	121.6	4.4	555	10	AW472190
43	121.2	4.4	794	11	BE406951
44	120.8	4.4	681	11	BG456652
45	120.4	4.3	665	11	BE406876

ALIGNMENTS

RESULT 1

BG343782

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BG343782 988 bp mRNA EST 29-MAY-2001
HVSMEG0006L20f Hordeum vulgare pre-anthesis spike EST library
HVCNAA0008 (white to yellow anther) Hordeum vulgare cDNA clone
HVSMEG0006L20f, mRNA sequence.
BG343782 GI:13156111
EST.
BG343782.1
barley.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 988)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu,
Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo,
T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and
Wood, T.
Development of a genetically and physically anchored EST resource
for barley genomics
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAACCCCTCAATAAGGG
High quality sequence stop: 845.
Location/Qualifiers
1. .988
/organism="Hordeum vulgare"
/cultivar="Morex"

BASE COUNT 243 a 223 c 291 g 211 t 20 others
ORIGIN

/db_xref="taxon:4513"
/clone="HVSMEg0006L20f"
/clone_lib="Hordeum vulgare pre-anthesis spike EST library
HVCDA0008 (white to yellow anther)"
/tissue_type="pre-anthesis spike"
/lab_host="SOLR"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"

Query Match 21.7%; Score 601; DB 11; Length 988;
Best Local Similarity 91.1%; Pred. No. 2e-52;
Matches 704; Conservative 0; Mismatches 59; Indels 10; Gaps 6;

Qy 1787 ccgagtcgagttacaagataaaattccgtggtggtggtgatttagttccagtttccc 1846
Db 1 CCGAGTCGAGCTACAAGGATAAATTCGTCGGATGGTGGATTAGTGTTCAGTTTCCC 60
Qy 1847 acagaataactcgaggtgagatattgtaatgccatcgagatttgaaccttgcggtc 1906
Db 61 ACAGAATAACTCGAGGTGCGATATATGTTAATGCCATCCAGATTTGAACCTTCCGGTC 120
Qy 1907 ttaacagctatgctatgcaatagtggtacagttctctgtagttcatggaactgggggcc 1966
Db 121 TTAATCAGCTATATGCTATGCAATATGGTACAGTTCCTGTTGTTTCATGGAATGGGGGCC 180
Qy 1967 tccgagacacagtcgagaccttcaaccttttgggtgcaaaaggagagaggtacaggg 2026
Db 181 TTCGAGACACGGTGAGACCTTCAACCCCTTTGGTGTGCAAAAGGAGAGGGGTACAGGGT 240
Qy 2027 gggggtttccacgtacacgtgacaaagattgttggtggttgcgaacgcgagtgcca 2086
Db 241 GGGGGTTCACACCTTCAACCGTGGAAATAATGTTGTGGCATTCGGAACCCGATTTCCA 300
Qy 2087 cattcaggagagacaaagcgttcctgggaggggtcctgaagcggaggtacgaagacc 2146
Db 301 CGTTTAGGGAGACAAAGCCCTCTGGGAGGGGCTCATGAAGCGAGGATGACGAAGACC 360
Qy 2147 atacgtgggaccatgccc-cgagacagtcagagcagagatttcgagtggtggttcgtgacc 2205
Db 361 ATACGTGGGACCATGCCGCCGAGAGTACGAGCAGATCTTCGAGTGGGCCCTTCGTGGACC 420
Qy 2206 aacctctacgtatgacgagggactgggaggtgccaaagtcgaagtcgagtccttcagctctg 2265
Db 421 AACCCCTACGTATGATGACTGGGACTGGGGAGGTGCGAAGCCGCCAGTCTCCTTGAGCTCTG 480
Qy 2266 aagacatcctcttcatccttcgcggcccgga---aggataccctgtacattgctgtgt 2322
Db 481 AAGACATGTTCTCATCTCTCGCGGCCCGGAAGGAGGATACCCCTGTACATTCGTTGT 540
Qy 2323 cctgtacagtagagtcgaatgcgctgttctgttctgttgcgaggttcgagacatat 2382
Db 541 CCTGCTACAGTAGAGTGCATATGCGCTTGTTCG--TTGGTTTCGCGGTTTCGAGAGTAGAC 599
Qy 2383 aecgactgtgctgcgaggggtgacagcttcgggtgacagacag-tcacagtttgggg 2441
Db 600 GACGTCGCAACCCCTGCGCCGGS---CAGCTTCGGGTGGATGACAGNNTACAGTTNTGGGG 656
Qy 2442 aataaggnaaggatgtgctgcaggatggttaacagcaagaagcaccactcagatggcagc-c 2500
Db 657 AATAAGGAAGGGAGTGTGTCAGGATGGNTGACAGAAAGACACACCTCAGTCGCGACCTC 716
Qy 2501 tctctgctcggtttacagctgaatcgaacaaactgtagctcttagcct 2553
Db 717 TCTCGGTGGGGGTACAGCTGAATCAGAAACCCACTGGTGTCTNCTTAGCCT 769

RESULT 2
AV832913/c

LOCUS AV832913 706 bp mRNA EST 22-JUN-2001
DEFINITION AV832913 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare top three leaves adult, heading stage Hordeum vulgare subsp. vulgare cDNA clone rbaal2b23, mRNA sequence.
ACCESSION AV832913 GI:14525002
VERSION EST.
KEYWORDS Hordeum vulgare subsp. vulgare.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 706)
AUTHORS Sato, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2001)
COMMENT Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kzsato@rib.okayama-u.ac.jp,
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct
submission;
database: http://www.shigen.nig.ac.jp/barley/barley.html.

FEATURES
source
1..706
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="rbaal2b23"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare top three leaves adult, heading
stage"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage" 2 others

BASE COUNT 174 a 213 c 162 g 155 t
ORIGIN

Query Match 20.4%; Score 564; DB 10; Length 706;
Best Local Similarity 91.9%; Pred. No. 1.2e-48;
Matches 651; Conservative 0; Mismatches 47; Indels 10; Gaps 5;

Qy 1914 gctatatgctatgcaatataggtacagttctgttagttcatggaactggggccctccaga 1973
Db 706 GCTATATGCTATGCAATATGTTGTGGCATTCGGAACCGGATTCGAGTTCGAGA 647
Qy 1974 cacagtcgagaccttcaaccttttgggtgcaaaaggagagaggtacaggtggcgctt 2033
Db 646 CACGGTGGAGACCTTCAACCCCTTTTGGTGCANAAAGAGAGAGGGGTACAGGGTGGCGTT 587
Qy 2034 ctcaccgtcaacgtggacaagatgttgggttcgcaacccgcgagtcgacattcag 2093
Db 586 CTCACCACTAACCGTGCAGAAAATGTTGTGGCATTCGGAACCGGATTCGAGTTCAG 527
Qy 2094 ggaacacacgcgttcctcgggaggggtcctgaagcagaggtacgaacacacacacgtg 2153
Db 526 GGAGCAACAGGCCCTCCCTGGGAGGGGCTCATGAGCGAGGATGACGAAGACCATACGTG 467
Qy 2154 ggacattgccc-ccgagcagtcacgagcagatcttcgagtgggcccttcgtggaccaacccta 2212
Db 466 GCACCATGCCGCCGAGCAGTACGAGCAGATCTTCGAGTGGGCCCTTCGTGGACCAACCTA 407
Qy 2213 cgtcatgtagacgggagactggggagggtccaaagtcgagtcctcttcagctctgaagacat 2272
Db 406 CGTCATGTAGACTGGGACTGGGGAGGTGCGAAGCGCCAGTCTCCTTTGAGCTCTGAAGACAT 347
Qy 2273 cctcttcactctccgcggcccgga---aggataccctgtacattgctgtcctgcta 2329
Db 346 GTTCTCTATCTTCCGCGGCCCGAAGAGGAGATACCCCTGTACATTGCGTTGCTCCTGCTA 287
Qy 2330 cagtagagtcgcaatgcgcctgctgtgttgggttcgcgggttcgagacaacatatgacggct 2389

```

Db 286 CAGTAGAGTCGAATCGCTTGCTTGC -TTGGTTCCGCGGTTCGAGAGTAGACGACGTC 228
QY 2390 gtgctgctggcgcggtgacagcttcgggtggagacagcttttgagggaataagga 2449
Db 227 GCAACGCTGCCGCGG---CAGCTTCGGGTGGATGACAGTTACAGTTTGGGGAATAAGGA 171
QY 2450 agggatgtgctgcaggtggttaacagcaaaagcaccactcagatggcagctctct--gt 2507
Db 170 AGGGAGGTGTCGAGGATGTTGACACGAACGACCACTCAGTCGGCAGGCTCTCTCGGT 111
QY 2508 ccgtgtacagctgaaatcagaacacaaactggtagctcttttagccttagtgattgtgaag 2567
Db 110 GGGTGTTACAGCTGAATCAGAAACCAACTGGTCTCTTTAGCCTTAGTGATGAAG 51
QY 2568 ttgttgcctctgtgtatgttctgttctttagcttagtgacaaatat 2615
Db 50 TTTGTCGCTCTGTGTATGTATCTTGTCTTGTCTTAGCTNATNATAATTT 3

RESULT 3
BE418354 988 bp mRNA EST 24-JUL-2000
LOCUS SCL031.D08R90531 ITEC SCL Wheat Leaf Library Triticum aestivum
DEFINITION cDNA clone SCL031.D08, mRNA sequence.
ACCESSION BE418354
VERSION BE418354.1 GI:9416200
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 988)
AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchlioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
TITLE International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
JOURNAL Unpublished (2000)
COMMENT Contact: Cloutier S
Cereal Research Centre, Agriculture & Agri-Foods Canada
Winnipeg MT CANADA
Tel: 204 983 2340
Fax: 204 983 4604
Email: scloutier@em.agr.ca
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

FEATURES
source
1..988
/organism="Triticum aestivum"
/cultivar="Thatcher Lrl"
/db_xref="taxon:4565"
/clone="SCL031.D08"
/tissue_lib="ITEC SCL Wheat Leaf Library"
/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/note="Vector: Lambda ZAP; 1.0 kbp average insert size."
BASE COUNT 227 a 199 c 244 g 247 t 71 others
ORIGIN

Query Match 17.8%; Score 494.4; DB 10; Length 988;
Best Local Similarity 96.9%; Pred. No. 1.1e-41;
Matches 564; Conservative 0; Mismatches 12; Indels 6; Gaps 6;

QY 1699 atggccattccagctcatgaggaggacgtgcaattgtctatctgtgattctgggag 1758
Db 7 AGGGCAATTCAGAGCTCATGAGGAGGAGGTGCAATTTGTCATGCTTGATCTGGGAT 66

```

```

QY 1759 ccaatttttaagctgtagatgagatctacagagtcaggttacagagataaattccgtgga 1818
Db 67 CCAATTTTGAAGGCTGGATGATATCTACCAGAGTCGAGTTACAGGATAAATTCGCTGGA 126
QY 1819 tgggttgattagttccagtttccacagataactcagattgcgatatattgtta 1878
Db 127 TGGTTGGATTAGTGTTCAGTTTCCACAGAATAACTGCAGGTTCGATATATCTTGA 186
QY 1879 atgcatcgagattgaaccttcggtcttaactaagctatatctatgcaatatggtaca 1938
Db 187 ATGCCATCGAGATTGAACCTTCGGTCTTTAATCAGCTATATGCTATGCAATATGTACA 246
QY 1939 -gttctctgtagtcatggaactgggacctccagacacacagctcgaaccttcaaccttt 1997
Db 247 CGTTCCTGTAGTTCATGGAACCTGGGGCCTCCGAGACACAGTCCAGACCTTCAACCCCTT 306
QY 1998 tgggtcaaaaggagaggggtacaggggtggcggtctctaccctcgaaccttgacaagat 2057
Db 307 TGGTGCAAAAGAGAGAGGGGTACAGGGTGGCGGTTCTCACCGCTAACCGTGGACAAGAT 366
QY 2058 gttgtgggcaattcgaaacgcgtagtgcacattcaggaggagacaaagccgtctctggaggg 2117
Db 367 GTTGTGGGCATTCGGAACCGCATGTCGACATTCAGGGAGACAAAGCCGTTCTGGGAGG 426
QY 2118 gtcattgaacggaggtacgacgaagaccata-cgtgggaccatgcc-cgagcagtcag 2175
Db 427 GCTNATNACCGAGGATGACGAAGAACCATACCTACCTGGGGACCATCGCCCGACGATCG 486
QY 2176 agcagatcttcgagtggtggccttcgtggaccaacctcactgca-tgtagacgggactggg 2234
Db 487 AGCAGATNTTTGAGTGGCGCTTCGTGGACCAACCTACNTCATTTGTAGACNGGAGCTGG 546
QY 2235 gag-gtccaaagtccagatctccc-ttcagctctgaagacatccc 2274
Db 547 GAGNGTCCAAAGTCGAGTCTCTTTGAGCTTTGAANACATCC 588

RESULT 4
BG607242/c 557 bp mRNA EST 17-APR-2001
LOCUS WHE2493_B06_D1LZS Triticum monococcum early reproductive apex cDNA
DEFINITION library Triticum monococcum cDNA clone WHE2493_B06_D1L, mRNA
sequence.
ACCESSION BG607242
VERSION BG607242.1 GI:13657225
KEYWORDS EST.
SOURCE Triticum monococcum.
ORGANISM Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 557)
AUTHORS Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia
,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,
Stamova,B. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat
genomes - Early reproductive apex cDNA library from Triticum
monococcum
JOURNAL Unpublished (2001)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
source
1..557
/organism="Triticum monococcum"
/cultivar="DV92"

```

/db_xref="taxon:4568"
 /clone_lib="WHE2493_B06_D11"
 /clone_lib="Triticum monococcum early reproductive apex
 cDNA library"
 /tissue_type="Early reproductive apex"
 /dev_stage="Seven week-old plants"
 /lab_host="E. coli XL0UR"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and
 poly(A) RNA were prepared from apex at double-ridge stage
 to terminal-spikelet stage during transition from
 vegetative state to flower state, a cDNA library was made,
 and the cDNA clones were in vivo excised at the
 University of California, Davis (V. Echenique, B. Stamova
 , J. Dubcovsky). Plasmid DNA preparations and DNA
 sequencing were performed in the OD Anderson lab (all
 other authors)."

BASE COUNT 140 a 161 c 118 g 138 t
 ORIGIN

Query Match 14.6%; Score 403.8; DB 11; Length 557;
 Best Local Similarity 88.9%; Pred. No. 2.1e-32;
 Matches 472; Conservative 0; Mismatches 52; Indels 7; Gaps 3;
 QY 2111 gggagggtctatgaagcgaggtacgacgaagaccatagctgggaccatgcc-ccgagc 2169
 DB 557 GGGAGGGGCTCATGACGAGGCGATGACGAAGACCATATGTGGACCATGCGCCGCGAGC 498
 QY 2170 agtacgagcagatcttcagtgggcctctgtgacacacccctacgtcatgtagacggga 2229
 DB 497 AGTACGACGAGATCTCGAATGGGCCCTCGTGGACCAACCATATGTATGTAGACAGG 438
 QY 2230 ctggggagggtccaaagtgcagctccctcaagctcgaagacatctctcatcttcctcgc 2289
 DB 437 CTGGGAGGTGCAAGCGCGAGTCTCTTGTAGCTCAGAAGACATGTTCTCATCTTCGCG 378
 QY 2290 ggcgggaagatgataccctgtacattgcgtgtctcgtctac-----agttagctcgcaat 2344
 DB 377 GCGCGGAAGGATACCCCTGTACATTCGCTGGAAACGCGGCTCTCTCAATAGTGGCAAT 318
 QY 2345 gcgcctctctgtgttgcgcgttccgagacacatagcagcgtctgtctgcgcgcgg 2404
 DB 317 GCGCCTGCTTGC-TTGGTTCGCCGGTTCGAGAGTAAATGATGGCTGTCTGCTCGCGCGG 259
 QY 2405 tgacgctctcgggtggacgacagttacagtttttggggaataagaaagggaagtgtcgtcag 2464
 DB 258 TGACAGCTTCGGGTGATGACAGTTACAGTTTTCGGGAATAAGAGGGGTGTGCGAGG 199
 QY 2465 gatggttaacagaaagaccacactcagatggcagcctctctgtcgtgttacagctgaaa 2524
 DB 198 AATGTTTAAACAGCAAGTTGACACTCAGCTGGCAGCCCTCCTGGTCCGTGTACAGCTGAAA 139
 QY 2525 tcagaaacaaactgtaactctttagcttagtattgtaagttgttgcctctctgtgt 2584
 DB 138 TCTGAACCAACTGGTGACTCTTTAGCCTTAGTGAATGTGAAGTTTGTTCCTCTCTGTGT 79
 QY 2585 atgtgtctgtctctagctgacaaaattttgacctgttgagaaattttat 2635
 DB 78 ATGTTGTCTTCTCTTACCTGACAAATTTTGTGACCTGTGTGGAGAAATTTTT 28

RESULT 5

AW678087 562 bp mRNA EST 19-JUL-2000
 LOCUS WSL_13_G10_b1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA,
 DEFINITION mRNA sequence.
 ACCESSION AW678087
 VERSION AW678087.1 GI:7551807
 KEYWORDS EST.
 SOURCE sorghum.
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 562)
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
 ,L.H.
 An EST database from Sorghum: water-stressed plants
 Unpublished (2000)
 Contact: Cordonnier-Pratt MM
 Department of Botany
 The University of Georgia
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 542 1805
 Email: mmp@plantuga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence start: 17
 High quality sequence stop: 553
 POLYA=NO.

FEATURES
 source

Location/Qualifiers
 1..562
 /organism="Sorghum bicolor"
 /db_xref="taxon:4568"
 /clone_lib="Water-stressed 1 (WS1)"
 /note="Organ: Mix of 5-week old plants on days 7 & 8 after
 water was withheld; Vector: Lambda zap; Site_1: XhoI;
 Site_2: EcoRI; The library was made from poly-A RNA in the
 cloning vector lambda ZAP II. Clones to be sequenced were
 prepared by mass excision."
 BASE COUNT 153 a 126 c 142 g 140 t
 ORIGIN

Query Match 14.2%; Score 393; DB 10; Length 562;
 Best Local Similarity 84.1%; Pred. No. 2.5e-31;
 Matches 466; Conservative 0; Mismatches 86; Indels 2; Gaps 2;
 QY 1660 ggaagactggtaccagaaaggcattgattctctataaaatggccattccagagctcatg 1719
 DB 10 GNAAGATTGGATTATCAGAAA-GCATTGATCTCATTTCACTTATCATACCATCTCATG 68
 QY 1720 agggagacgtgaattgtcattgcttgatctgggagatcccaatttttgaggcgtgatg 1779
 DB 69 CGGACACGCTTCAATTTGTTCATCTGCTGGATCTGGTACCACAGAGCTCGAAGACTGGATG 128
 QY 1780 agactaccgagtcagattacaaggataaattccgtgagtggttgatttagtctcca 1839
 DB 129 AGATCTACAGATCGGACTTCAGAGTAATAATTCGTGGATGGGTGGATTAGTGTCCA 188
 QY 1840 gttcccccacagaataactgcaggttgcgatatattgttaatgccatcgagattgaaacct 1899
 DB 189 GTTTCACCGAATAACTGCGCGCTGGATATATTGTTAATGCCATCCAGATTCGAACCT 248
 QY 1900 tgcggttttaacagctatattgctatgtaataatgtaacagttcctgtagttcagaaact 1959
 DB 249 TGTGGTCTCAATCAGCTATATGCTATGATGATGGCAGTTCCTGTGTGCCATGCAACT 308
 QY 1960 gggggcctccgagacacagtcgagaccttcaaccttttgggtcaaaagagagagaggt 2019
 DB 309 GGGGGCCTTAGAGATACTGTGGAGAACTTCAACCCCTTCGGTGAGATGGAGAACAGGGT 368
 QY 2020 acaggggtggcggttctcaccgctaaacctggaagaagtgtgtgtgggcatctgcgaaccgcg 2079
 DB 369 ACAGGGTGGCATTTCCGACCCCTTAACACGGAATAATGTTGTGACATTTGCCAATGCA 428
 QY 2080 atgtcgacattcaggagacacagccctcctggagggtcctatgtaagcagcagatgacg 2139
 DB 429 ATTTTCAGATACAAAGAGACAAATCTCTCTGGGAAGGGCTCATGAAGAGGGCATGTCA 488
 QY 2140 aaagaccatcagtggggaccatgcc-ccgagcagtcacgagcagatcttcgagtgggccttc 2198

Db 489 AAGACTTCACGTGGGACCATCCGCTGAACAATAACGAACAAATCTCCAGTGGCCCTTC 548

QY 2199 gtggaccaccccta 2212 22-JUN-2001 EST

Db 549 ATCGATCGACCTA 562

RESULT 6

AV832689 454 bp mRNA 22-JUN-2001 EST

LOCUS AV832689 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare top three leaves adult, heading stage Hordeum vulgare subsp. vulgare cDNA clone baal2b23, mRNA sequence.

DEFINITION AV832689.1 GI:14524778

ACCESSION AV832689

VERSION AV832689

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare.

ORGANISM Hordeum vulgare subsp. vulgare.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.

AUTHORS Sato, K.

TITLE Barley EST sequencing project in NIG and Okayama Univ

JOURNAL Unpublished (2001)

COMMENT Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazsato@rib.okayama-u.ac.jp
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct submission;
database: http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES

source 1. .454

BASE COUNT 112 a 84 c 126 g 130 t 2 others

ORIGIN

Query Match 14.0%; Score 388; DB 10; Length 454;
Best Local Similarity 95.1%; Pred. No. 8.8e-31;
Matches 411; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 592 gaaggattgctgagattccatcgacagcataattgtgctgcaagtgcaggattct 651

Db 24 GGAGGATTGCTGAGGATTCC-TCGACACCATAGTTGTGGCTGCAAGTCAGCAGGATTCT 82

QY 652 gagatcatgatgcaaatgacaaacctcaagctaaagtacacgtacgtcgtgtgtgtg 711

Db 83 GAGATCATGGATGCCAACGACCAACCTCTAGCTAAAGTTACACGATGATCGTGTGTG 142

QY 712 actggtgaagctgctccttatcaaaagtcagggggttggagatgttggttcggtta 771

Db 143 ACTGTTGAAGCTGCTCTTATCAAGTCAGGGGGCTGGGAGATGTTTGGTTCGTTG 202

QY 772 ccaattgcttctgctgctggtgcacgcagtgatggtgttaatgccaaagataactaaat 831

Db 203 CCAATTGCTCTTCTGCTGCTGGTGCACCGTGTGTGTTGTTAAATGCCGAGATCTTAAAT 262

QY 832 gggctccttgaLaaaaactatgcaaaagcattatcacactgcgaagcacattcaagattcca 891

Db 263 GGGACCTCTGATAAAACTATGCAAAAGGCATTATACACCGGGAAGCACATTAAAGATTCCA 322

QY 892 tgctttggggatcacatgaagtgaaccttttttcatgtagtatagagacaacgctgattgg 951

Db 323 TGCTTCGGGGATCACATGAAGTCACCTTTTTCATGAGTATAGACAACGTCGATTTGG 382

QY 952 gtgtttgcatcatccgtcatatcacagaccaggagtttatatgagataattttggt 1011

Db 383 GTGTTTGTTCATCATCATCATATAGACACCAAGTGTGTATGGAGATAATTTTGGT 442

QY 1012 gctttggtgat 1023

Db 443 GCTTTTGGTCAT 454

RESULT 7

BF482660 408 bp mRNA EST 06-DEC-2000

LOCUS WHE2301-2304_E13_E13S Wheat pre-anthesis spike cDNA library

DEFINITION Triticum aestivum cDNA clone WHE2301-2304_E13_E13, mRNA sequence.

ACCESSION BF482660

VERSION BF482660.1 GI:11565884

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat genomes - pre-anthesis spike cDNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: StrataGene SK primer.

FEATURES

Location/Qualifiers

source 1. 408

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE2301-2304_E13_E13"

/clone_lib="Wheat pre-anthesis spike cDNA library"

/tissue_type="Spike before anthesis"

/dev_stage="Adult plant"

/lab_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 108 a 76 c 99 g 125 t

ORIGIN

Query Match 13.7%; Score 380.8; DB 11; Length 408;
Best Local Similarity 95.8%; Pred. No. 4.9e-30;
Matches 391; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1558 tattctgcatgacctctctggaaggccaaatgaaagctgaattgcagaaggagttg 1617

Db	1	TATTCTGTCGATGACCTCTCTGSAAGGCCAAATGTAAGCTGAATTGCACAGGAGCTG	60
QY	1618	ggtttacattgaaggagatgtctctgtgattggtctttattggaagactggattaccag	1677
Db	61	GGTTTTACCTGTAAGGAGATGTCTCTGATGGCTTTATTGGAAGACTGGATTACCAC	120
QY	1678	aaaggcattgatctcattaaatggccattccagagctcatcagggagagcgtgcaattt	1737
Db	121	AAAGGATTGATCTCATTAATTAATGGCCATTCACAGAACTCATGAGGAGGACGTACACTTT	180
QY	1738	gtcatgcttgatctggggatccaatttttgaagcgtgatgatcaccagtcgagtc	1797
Db	181	GTATGCTTGATCTGGGGATCCAAATTTTGAAGGCTGGATGAGATCTACCGAGTCTAGT	240
QY	1798	tacaagataaaatccgtggtggtggttagttggttcagtttccccacagaataact	1857
Db	241	TACAAGGATAAATTCGCTGGATGGTGGATTTACTGTCCCAATTTCCACACAAATACT	300
QY	1858	gcaggttcgatataattgttaagccatcgagatttgaaccttcgcttcaatcagcta	1917
Db	301	GCATGTTGCCGATATATTGTTAATGCCATCCATGTTTGAACCTTGTGTCITTAATCAACTA	360
QY	1918	tatctatgcaaatggtacagttccctgtagttcattgagaaactgggggc	1965
Db	361	TATGCTATGCAATATGTACAGTTCTCTGTAGTTTCATGGAACCTTGGGGGC	408
RESULT	8		
LOCUS	BF113316	596 bp mRNA	18-MAY-2001
DEFINITION	EST440906 tomato breaker fruit Lycopersicon esculentum cDNA clone CLEG44E24 5' sequence, mRNA sequence.		
ACCESSION	BF113316		
VERSION	BF113316.1	GI:10943006	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		
REFERENCE	1 (bases 1 to 596)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley S.D.		
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: CUGI Clemson University Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html .		
FEATURES	Location/Qualifiers		
source	1..596		
	/organism="Lycopersicon esculentum"		
	/cultivar="TA496"		
	/db_xref="taxon:4081"		
	/clone="CLEG44E24"		
	/tissue_type="tomato breaker fruit"		
	/tissue_type="pericarp"		
	/dev_stage="breaker"		
	/lab_host="SOLR"		
	/note="Vector: pBluescriptSKmQuadapt; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."		
BASE COUNT	167 a 120 c 149 g 160 t		
ORIGIN			
Query Match	11.8%; Score 325.6; DB 11; Length 596;		

Best Local Similarity	71.6%;	Pred. No. 1.6e-24;	
Matches	427;	Conservative	0; Mismatches 169; Indels 0; Gaps 0;
QY	1156	cttctgtgtgcaaaatatagaccatacagtggtttattacagagattccctccagcacccttatt	1215
Db	1	CTTTTAGCAGCCAAAGTATCGTCTTATGTTTAAAGGATGCTCGTAGTATTGTCGCA	60
QY	1216	atacataatttagcacatcaggtgtgagcctgcaagtacatatccctgatctctgggattg	1275
Db	61	ATACACAACTTGCACATCAGGAGCTGGAGCCTGCAGCAACCTACAATAATTTGGGATTG	120
QY	1276	cctcctgaatggdtagggagctttagaatgggttattccagaatgggcaaggaggccatccc	1335
Db	121	CCTCCTCAATGCTATGGAGCACTTGAATGATATTTCACATGGCAAGGCCCATCGG	180
QY	1336	cttgacaagggtgagcgagtttaactttttgaaaggcgagttgtgacagcagatcgagatt	1395
Db	181	CTTGACACTGGGTGAACAGTGAATGTTTCAAAAGGGCAATCTCAGTTGCTGATCGGATA	240
QY	1396	gtgaccgtcagtcaggggttattcatcagggaggttcacaactgctgaaagggtgacagggccctc	1455
Db	241	CTGACAGTTAGCCAGGGATACTCATGGGAAATAACAACCTCTGAAGGGGATATGGGTTA	300
QY	1456	aatgagctcttaagctcccccgaagaaagtgtattgaatgggaattgtaaatggaattgacatt	1515
Db	301	CATGAGCTGCTGAGCAGTAGACAGTCAGTCTTAAATGGAAATTAATAAGTAAGATGTT	360
QY	1516	aatgattggaaccccccacacagacaagtgtctctccctcatcattatctctcgatgacctc	1575
Db	361	AATGATTGGNACCCGTCGACAGATGAGCATATATGCTTCGCATTACTCCATCAATGACCTC	420
QY	1576	tctggaagagcccaaaagttaaagctgaatgcagaaagagttgggtttacctgtaagggag	1635
Db	421	TCTGGAAGGCTCAGTGCAGACTGATCTGCAAAAGAACTGGGCCCTTCCAATTCACCT	480
QY	1636	gatgttcctctgattggtcttattggaagactggtattaccagaaagcattgtatctcatt	1695
Db	481	GATTGTCGGCTGATGGATTTATTGGNAGCTGGACTACCAGAAAGGTGTTGACATAATC	540
QY	1696	aaaatggccattccagagctcatcagggaggacgtgcaattttcatgcttgatc	1751
Db	541	CTGTCAGCAATTCAGAACTTTTGCAGAGGATGTCCTCAATTTGTAATGCTTGGATC	596
RESULT	9		
LOCUS	BE434991	588 bp mRNA	EST 18-MAY-2001
DEFINITION	EST406069 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone CLEG24J21, mRNA sequence.		
ACCESSION	BE434991		
VERSION	BE434991.1	GI:9432834	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		
REFERENCE	1 (bases 1 to 588)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley S.D.		
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence.		
FEATURES	Location/Qualifiers		
source	1..588		


```
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db_xref="taxon:4081"  
/clone="cLEG24721"  
/clone_lib="tomato breaker fruit, T  
/tissue_type="pericarp"  
/dev_stage="breaker"  
/lab_host="SOLR"
```

```

/note="Vector: pBluescriptSMCquadapt; site_1: EcoRI;
site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
115
116

```

BASE COUNT
ORIGIN

```
Query Match      11.6%; Score 320.2; DB 10; Length 588;
Best Local Similarity 71.9%;
Pred. No. 5.5e-24;
Matches 418; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
```

Qy 1211 ttgttatacataatttagcacatcagggttgtgagcctgcgaagtacatatctgatctgg 12
| | | | | | | | | | | | | | | | | | | | |
Db 4 TCGCAATACACAACATTCCACATCAGGGACTGAGCCTCGACGAACCTTACATAAATTTGG 63

Qy 1271 gattgctcctgaatggtatggagctttaagatgggtattccagaatggcgaagagc 13

Db 64 GATTGCCCTCCTCAATGGTATGGAGCACTTCAATGGATATTTCCCAATGGCGAAGGCC 12

Qy 1331 atgcccttgacaagggtgaggcaggttaactttttgaaaggagcagttgtgacagcagatc 1390

DB 124 ATGCGCTTGACACTGGTGAACAGTGAATGTTTGAAGGGCAATCTCAGTTGCTGATC 183

QY 1391 ggattgtgaccgtcagtcagggttattcatatggaaggtcacactgctaaaggttgcacag 1450

Db
184 GGATACCTACAGTTAGCCAGGATCATCGGAAATACAACTCCTGAAGGGGATATG 243

Qy	1451	gcctcaatgagcctttaaagtccccgaanaagtgtattgaatggaattgtataatggaaattg	1510
Db	244	GGCTACATGAGCTGCTGAGCAGTGTAGACAGTCAGCTCTTTCTTCCCACTAATCGATTCTT	303

Oy 1511 acattaatgattggaaacccaccacagacaagaagtgtctccotcatcattattctgtcgatg 1570
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
-----TGGGAGGATGACGAGCTTCATTGAAGAATTAATAATGGAATAG 303

304 ATGTTAATGATTGGAACCCGTCGACAGATGAGCATATGTGTCGATTACTCCATCAATG 363

db

364 ACCTCTCTGGAAGGCTCAGTGCACACTGATCTGCAAAGGAACATGGGGCCCTCCAAATTC 423

1631 gggaggatgtctctctgattggctttattggaagactgattaccagaaaggcattgac 1690

424 GACCTGATTGTCGGCTGATTGGATTATTGGAGGGCTGGACTACCAAGAAAGTGTTGACA 483

Qy 1691 tcattaaaaatggccattccagagagctcatgagggagggagcgtgcaattttgtcatccttgaat 1750

484 TAATCCTGTCAGCAATTCAGAACTTTGCAGAAGGATGTCCAATTGTATGCTTGGAT 543

Qy 1751 ctgggagatcccaattttgaaggctggatgagatctaccgag 1791
||||| | | ||||| ||||| ||| |||
Db 544 CTGGTGAGAAACAATATCAAGACTCCATCAGACATACACAC 584

RESULT	10
LOCUS	AU075475
DEFINITION	396 bp mRNA EST 14-JUN-1999
ACCESSION	AU075475
VERSION	AU075475.1
KEYWORDS	GI:5056096
SOURCE	EST.
ORGANISM	Oryza sativa.
	Oryza sativa
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

National Institute of
Rice Genome Research
305-8602, Japan
Tel: 81-298-38-7441

Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>
PROJECT = 'BCD'

FEATURES SOURCE

```
1. 350
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="E60759_1A"
/clone_lib="Rice cDNA fr
meristem"
```

[illegible]

Query Match 10.6%; Score 294.2; DB 10; Length 396;
Best Local Similarity 85.8%; Pred. No. 2.8e-21;
Matches 338; Conservative 0; Mismatches 54; Indels 2; Gaps 1;

Qy 1597 gctgaattgcagaagagittgggtttacctgtaaaggaggatgttctctctgattgcttt 1656
|||||
Db 2 GCTGAATTGCAGAAGAGCTGGGTTTACCTATAAGCCCGCATGCCCTCATTTGGGCTTT 61

Qy 1657 attggaagactggattaccagaaaggcattgatctcattaaaaatggccattccagagctc 1716
|||||
Db 62 ATTGGAAGATTGGACATATCAAAAGGCATTGATCTAATAAATCGCATTCAGATCTC 121
|||||

[illegible]

Qy 1777 atgagatctaccgagctcgaggtttacaaggataaaattccgtgatgggttgagttagtggtt 1836
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 ATGAGATCCACAGAATCAGGGTACAGGGATAAATTCTCGATCGGTGGATTGATGTTT 241

Oy 1837 ccagttctccacagaaactgcaggttcgatatatttgtaatagcgatcgagatttga 1896
|||||
|||||
Db 242 CCAGTTTCCACCGAATAACTCAGGTTGCATATATTGTGTGATGCCATCCAGATTCGA 301

Qy	1897	cttgcgggtcttaatcagctatagctatgcaatatggtacagttccctgtagttcatgg-	1955
Db	302	CCTTTGGCGCTCAATCAGCTATATGCTATGCAATATGGTACAGTCCCTGNTGTTCATGGG	361

Qy	1956	-aacTgggggcctcgagacacagtcgagacctt	1988
Db	362	AACtgAGGGcCTCAGATAcAgTGGAGAATTt	395

RESULT 11
BG050773

LOCUS	BG050773	401 bp	EST	25-JAN-2001
DEFINITION	FW1_70_D03.b1_A003 Floral-Induced Meristem 1 (FW1) Sorghum propinquum cDNA, mRNA sequence.			
ACCESSION	BC050773			

RECORD	VERSION	KEYWORDS	SOURCE	ORGANISM
B050773	B050773.1	EST.	Sorghum propinquum.	Sorghum propinquum.
				Fukaryota: Viridiplantae

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE 1 (bases 1 to 401)
AUTHORS Cordonnier-Pratt M.M., Gingle A., Sudman M., Marsala C. and Pratt L.H.
TITLE An EST database from Sorghum: floral-induced meristems
JOURNAL Unpublished (2000)
COMMENT Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 361
POLYA-No. Location/Qualifiers
1. .401

FEATURES
source
1. .401
/organism="Sorghum propinquum"
/db_xref="taxon:132711"
/clone_lib="Floral-Induced Meristem 1 (FMI)"
/note="Organ: Floral-Induced Meristems; Vector: pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."
102 a 72 c 104 g 123 t

BASE COUNT 102 a 72 c 104 g 123 t
ORIGIN
Query Match 10.0%; Score 276; DB 11; Length 401;
Best Local Similarity 90.7%; Pred. No. 1.9e-19;
Matches 294; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1262 ctgattcgggattgctcctgaatggatggagcttttagaattgggtatttcacagaatggg 1321
Db 1 CTGACCTTGGTGGCCACCTGAATGGAGCTGGAGCTCTGGAGTGGGTATTTCCTGAATGGG 60
Qy 1322 caaggaggatgccttgacagggtgaggcaggttaacttttcaaaaggagcagttgtga 1381
Db 61 CAAGGAGGCATGCCCTTGACAAAGGTGAGGCAGTTAAATTTTGAAGGTCAGTTGTGA 120
Qy 1382 cagcagatcggttgaccgacagtcagggttattcattgaggtgcacaaactgctgaag 1441
Db 121 CAGCAGATCGAATGTGACTGTCACTAAGGTTATTTCATGGGAGTCAACAAGTCTGAAG 180
Qy 1442 gtggacagggcctcaatgagctcttaagctcccgcaaaagtgtatgtaattgaaatgtaa 1501
Db 181 GTGGACAGGGCCTCAATGAGCTTTAAGCTCCGAAAGAGTGTATTAACGGAATGTAA 240
Qy 1502 atggaattgacattaatgattggaacccaccacagacaagtgctcctcatcattatt 1561
Db 241 ATGGGAATTGACATTAATGATTGGAACCTTGGACGGACATATGTATCCCTTGTCTCATATT 300
Qy 1562 ctgtcgatgacacctctctggaagg 1585
Db 301 CTGTTGATGACCTCTCTGGAAG 324

RESULT 12
AW932199 564 bp mRNA EST 18-MAY-2001
LOCUS EST358042 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION CDNA clone cLEF47J7 5', mRNA sequence.

AW932199
AW932199.1 GI:8107600
EST
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 564)
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,
Upson, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.,
Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. .564

FEATURES
source
1. .564
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_lib="cLEF47J7"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; cLEF - Fruit were tagged at the lcm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
150 a 115 c 140 g 159 t

BASE COUNT 150 a 115 c 140 g 159 t
ORIGIN
Query Match 9.0%; Score 250.6; DB 10; Length 564;
Best Local Similarity 67.7%; Pred. No. 6e-17;
Matches 381; Conservative 0; Mismatches 154; Indels 28; Gaps 1;
Qy 1136 atgccagccttgccagtcctcttctgcacaaatagaccatcggtgtttacagag 1195
Db 1 ATGCTTCCCTGGTTCCTTTACTTTTACGACGCCAAGTATCGTCTTATGGTGTTCACAAGG 60
Qy 1196 attcccgagcaccctgtttatatacataatttagcacatc----- 1234
Db 61 ATGCTCTGATGTTGTCGAATACACAACATTCACATCAGATCGTCTCTTTTGTATAT 120
Qy 1235 -----agggtgtggagcctgcgaagtacatatctgattctgggattgctcctcgaatgg 1287
Db 121 TTTCATATAGGAGTGGAGCTGCGACCAACCTACAAATAATTTGGGATTTGCCCTCCCAATGG 180
Qy 1288 tatggagctttagaatagggtatttcccaagaatggcagaggagcagccttgacaaagggt 1347
Db 181 TATGGACACATTTGAATGGATATTTCCACATGGGCAAGGCCCAATGCGCTTGACACTGGT 240
Qy 1348 gaggcagtaacttttgaaggagcaggtgtgacagcagatcgaggttgagcgcagtt 1407
Db 241 GAAACAGTGAATGTTTGAAGGGGCAATCTCAGTTGCTGATCGGATACGATACAGTGTAGC 300
Qy 1408 cagggttattcattggagggttcacaactgctgaaggtggcagggccctcaatgaactctta 1467
Db 301 CAGGGATACTCATGGGAAATAACAACCTCCTGAAGGGGATATGGCTACATGAGCTGCTG 360
Qy 1468 agctcccccacaaagtgtattgaatgggaattgtaaatgaattgacattgaattggagac 1527
Db 361 AGCAGTAGACAGTCAGTCTTCTTAATGGGAATTAATAATGAATAGATGTTAATGATTGGAAC 420

COMMENT

Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/.

FEATURES

source
Location/Qualifiers
1..335
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="E10921_1A"
/clone_lib="Rice panicle at ripening stage"
/dev_stage="ripening stage"
/note="Organ: panicle; Rice cDNA from panicle at ripening stage"

BASE COUNT 80 a 66 c 98 g 86 t 5 others
ORIGIN

Query Match

Best Local Similarity 7.8%; Score 216.4; DB 11; Length 335;
Matches 281; Conservative 0; Mismatches 36; Indels 6; Gaps 4;

```

Qy 1125 gaacgattggcatgccagccttgccagtccttctgtgcaaaatatagaccacagg 1184
Db 1 GAATGATTGGCATGCCAGCTCTGTGCCAGTCCTTCTGTGCAAAATATAGACCATATGG 60

Qy 1185 tgtttacagagattcccgccagacccttgttatacataatttagcacatcagggtgtgga 1244
Db 61 TGTTTACAGGAGTGCCCGCAGTGTCTTGTGCATACATAATCTAGCACATCAGGGTGTGGA 120

Qy 1245 gcttgaagtacatatctctatctgtggattgacctcctgaatggtatggagctttagaatg 1304
Db 121 CCCTGCCAGTACATATCTGACCTGGGATTGCCACCCTGAATGATGGAGCATTAGAATN 180

Qy 1305 ggtattccagaatgggcaaggagcgatgcccttgacaagggtgagcgaggttaactttt 1364
Db 181 GGTGTTTCCAGAGTGGGCAAGCGGCATG--CCPTGACAAGGGTGTAGGCAGTCAATTTTTT 239

Qy 1365 gaagagagcag-ttgtgacagcagatcgg-attgtgacctcagtcaggggtta---ttca 1419
Db 240 AAAAGGCGCAGTTTGTGNACAGCAGATCGGAATTGTNACTGTCAGCCAGGGGNAATTTCAT 299

Qy 1420 tggagggtcacactgctaagg 1442
Db 300 GGGGNGGTCACAACTGCTGAAG 322

```

Search completed: March 28, 2002, 13:48:18
Job time: 5992 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 15:16:00 ; Search time 238.29 Seconds
(without alignments)
880.896 Million cell updates/sec

Title: US-09-674-824-2
Perfect score: 4044
Sequence: 1 MAATGVGAGCLAPSRLRAD.....SDGSLSVRTAETRNLVTL 756

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues
Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents,AA,Main.*
1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US10_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US11_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US12_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US13_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US14_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US15_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US16_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US17_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US18_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US19_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US20_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US21_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US22_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US23_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US24_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US25_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US26_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US27_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US28_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US29_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	2557.5	63.2	583	18 US-09-402-254-53 Sequence 53, Appl
2	2557.5	63.2	583	20 US-09-625-406-13 Sequence 13, Appl
3	2390.5	59.1	539	20 US-09-625-406-21 Sequence 21, Appl
4	2189	54.1	435	24 US-60-312-544-9109 Sequence 9109, Ap
5	1992	49.3	641	20 US-09-606-304-10 Sequence 10, Appl
6	1718.5	42.5	459	20 US-09-606-304-4 Sequence 4, Appl
7	1374	34.0	303	24 US-60-324-109-29423 Sequence 29423, A
8	1337	33.1	297	24 US-60-312-544-5760 Sequence 5760, Ap
9	1276	31.6	535	24 US-60-324-109-21228 Sequence 21228, A

10	1225.5	30.3	670	18 US-09-402-254-51 Sequence 51, Appl
11	1215	30.0	767	20 US-09-606-304-8 Sequence 8, Appl
12	1209.5	29.9	804	18 US-09-402-254-49 Sequence 49, Appl
13	1205.5	29.8	698	20 US-09-623-406-11 Sequence 11, Appl
14	1196.5	29.6	801	17 US-09-388-743-26 Sequence 26, Appl
15	1184	29.3	558	20 US-09-606-304-6 Sequence 6, Appl
16	1142.5	28.3	650	17 US-09-388-743-6 Sequence 6, Appl
17	1068	26.4	669	20 US-09-625-406-9 Sequence 9, Appl
18	1055.5	26.1	440	24 US-60-324-109-17143 Sequence 17143, A
19	973.5	24.1	385	24 US-60-312-544-10029 Sequence 10029, A
20	891.5	22.0	636	24 US-60-288-315-4 Sequence 4, Appl
21	881	21.8	637	18 US-09-402-254-47 Sequence 47, Appl
22	880	21.8	477	19 US-09-583-110-3812 Sequence 3812, Ap
23	880	21.8	484	15 US-09-107-433-4468 Sequence 4468, Ap
24	877	21.7	609	20 US-09-625-406-7 Sequence 7, Appl
25	875.5	21.6	616	17 US-09-388-743-14 Sequence 14, Appl
26	870	21.5	604	21 US-09-731-166-4 Sequence 4, Appl
27	865	21.4	533	20 US-09-625-406-5 Sequence 5, Appl
28	860	21.3	614	17 US-09-388-743-18 Sequence 18, Appl
29	852.5	21.1	600	17 US-09-388-743-22 Sequence 22, Appl
30	840	20.8	342	24 US-60-312-544-10086 Sequence 10086, A
31	828.5	20.5	615	17 US-09-388-743-2 Sequence 2, Appl
32	826	20.5	599	24 US-60-288-315-5 Sequence 5, Appl
33	824	20.4	609	24 US-60-324-109-18806 Sequence 2, Appl
34	821	20.3	609	24 US-60-288-315-2 Sequence 18806, A
35	796.5	19.7	479	23 US-09-902-540-14647 Sequence 14647, A
36	696	17.2	480	21 US-09-739-449-8399 Sequence 8399, Ap
37	696	17.2	480	22 US-09-803-110-8399 Sequence 8399, Ap
38	691.5	17.1	548	16 US-09-252-691-8864 Sequence 8864, Ap
39	691.5	17.1	548	16 US-09-252-691-8864 Sequence 8864, Ap
40	681.5	16.9	511	18 US-09-489-039A-11033 Sequence 11033, A
41	678.5	16.8	477	3 US-07-731-226-2 Sequence 2, Appl
42	678.5	16.8	477	4 US-08-016-881A-12 Sequence 12, Appl
43	678.5	16.8	477	8 US-08-484-434A-12 Sequence 12, Appl
44	678.5	16.8	477	17 US-09-384-361-12 Sequence 12, Appl
45	650	16.1	236	24 US-60-324-109-25214 Sequence 25214, A

ALIGNMENTS

RESULT 1
US-09-402-254-53
; Sequence 53, Application US/09402254
; GENERAL INFORMATION:
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; FILE OF INVENTION: HOSTS
; FILE REFERENCE: 2461-52
; CURRENT APPLICATION NUMBER: US/09/402,254
; CURRENT FILING DATE: 1993-10-01
; EARLIER APPLICATION NUMBER: PCT/US98/06660
; EARLIER FILING DATE: 1998-04-03
; EARLIER APPLICATION NUMBER: 60/042,939
; EARLIER FILING DATE: 1997-04-04
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-402-254-53

Query Match 63.2%; Score 2557.5; DB 18; Length 583;
Best Local Similarity 86.7%; Pred. No. 1.5e-213;
Matches 476; Conservative 25; Mismatches 45; Indels 3; Gaps 2;
QY 45 VAEISREGPAARPAQOOLAPPLVPGGLAPPPAPAPASPAOTPPPLPDAGVGLAPDLLL 104
DB 2 VAEISREGPAARPAQOOLAPPLVPGGLAPPPAPAPASPAOTPPPLPDAGVGLAPDLLL 104

QY 105 EGAEDSIDSIIVASEQDSEIMDANEQOAKVTRISIVFTGEAAYAKSGGLGDCVCSL 164
DB 59 EGAEGSIDNTVVASEQDSEIVVGEQAKRAKVTQSVFVTGEASPYAKSGGLGDCVCSL 118
QY 165 PIALAARGHRVVMVMPRYLNGSDKNYAKALYAKIKIPCFGGSHEVTFHEYRDNDV 224
DB 119 PVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHEYRDNDV 178
QY 225 VFVDHPSYHRPGSLYGNFAGDNOFRYTLCCYAAACEAPLIILELGGYIYGQNCMFVND 284
DB 179 VFVDHPSYHRPGSLYGNFAGDNOFRYTLCCYAAACEAPLIILELGGYIYGQNCMFVND 238
QY 285 WHASLVVLLAAKRYPGYVVRDSTRILVTHLAHQGVPEASTYVDGLGPPWYGALEWVF 344
DB 239 WHASLVVLLAAKRYPGYVVRDSTRILVTHLAHQGVPEASTYVDGLGPPWYGALEWVF 298
QY 345 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 404
DB 299 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 358
QY 405 GIVNGIDINDWNPATDKCLPHHYSDVDDLSGKAKCKALQKELGLPIRPDVLGFIGRLD 464
DB 359 GIVNGIDINDWNPATDKCIPCHYSVDDLSGKAKCKALQKELGLPIRPDVLGFIGRLD 418
QY 465 YQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHR 524
DB 419 YQKGDILQIILPDLMRDQFVNLGSDPPELEDMMRSTESIFKDKFRGWGFSVPVSHR 478
QY 525 ITAGCDILLMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGAKGEEGTGWA 584
DB 479 ITAGCDILLMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGGEGQGTGWA 538
QY 585 FSPLTVDKM 593
DB 539 FAPLTENN 547

RESULT 2

US-09-625-406-13
; Sequence 13, Application US/09625406
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/625,406
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/941,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-625-406-13

Query Match 63.2%; Score 2557.5; DB 20; Length 583;
Best Local Similarity 86.7%; Pred. NO. 1.5e-213;
Matches 476; Conservative 25; Mismatches 45; Indels 3; Gaps 2;

QY 45 VAELESREGPAAPPAQOQQLAPPLVGFGLAPPPAPPAQSPAPTPQPLPDAGVCELAPDLLL 104
DB 2 VAELESREGPAAPPLAPPLVGFGLA-PPAEPTEGPASTPPPPVPPDAGLGDLG--LEP 58
QY 105 EGAEDSIDSIIVASEQDSEIMDANEQOAKVTRISIVFTGEAAYAKSGGLGDCVCSL 164
DB 59 EGAEGSIDNTVVASEQDSEIVVGEQAKRAKVTQSVFVTGEASPYAKSGGLGDCVCSL 118
QY 165 PIALAARGHRVVMVMPRYLNGSDKNYAKALYAKIKIPCFGGSHEVTFHEYRDNDV 224
DB 119 PVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHEYRDNDV 178
QY 225 VFVDHPSYHRPGSLYGNFAGDNOFRYTLCCYAAACEAPLIILELGGYIYGQNCMFVND 284
DB 179 VFVDHPSYHRPGSLYGNFAGDNOFRYTLCCYAAACEAPLIILELGGYIYGQNCMFVND 238
QY 285 WHASLVVLLAAKRYPGYVVRDSTRILVTHLAHQGVPEASTYVDGLGPPWYGALEWVF 344
DB 239 WHASLVVLLAAKRYPGYVVRDSTRILVTHLAHQGVPEASTYVDGLGPPWYGALEWVF 298
QY 345 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 404
DB 299 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 358
QY 405 GIVNGIDINDWNPATDKCLPHHYSDVDDLSGKAKCKALQKELGLPIRPDVLGFIGRLD 464
DB 359 GIVNGIDINDWNPATDKCIPCHYSVDDLSGKAKCKALQKELGLPIRPDVLGFIGRLD 418
QY 465 YQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHR 524
DB 419 YQKGDILQIILPDLMRDQFVNLGSDPPELEDMMRSTESIFKDKFRGWGFSVPVSHR 478
QY 525 ITAGCDILLMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGAKGEEGTGWA 584
DB 479 ITAGCDILLMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGGEGQGTGWA 538
QY 585 FSPLTVDKM 593
DB 539 FAPLTENN 547

RESULT 3
US-09-625-406-21
; Sequence 21, Application US/09625406
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/625,406
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/941,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:

APPLICATION NUMBER: US/09/625,406
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/941,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 21:
LENGTH: 539 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-625-406-21

Query Match 59.1%; Score 2390.5; DB 20; Length 539;
Best Local Similarity 88.4%; Pred. No. 4.9e-199;
Matches 444; Conservative 23; Mismatches 32; Indels 3; Gaps 2;
QY 95 VGELA-PDLLL--EGIAEDSIDSIIVAAEQDSEIMDANQPOAKVTRTSIVFTVGEAAPY 151
Db 2 VAELSREDGLLEPAGIEGSDINTVVVASEQDSEIVGVKEQAKVTSIVFTVGEASPY 61
QY 152 AKSGGLGDCVSLPILAAARGHRVWVMPRYLNGSSDKNYAKALYTAKHKIPCFGGSHE 211
Db 62 AKSGGLGDCVSLPILAAARGHRVWVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHE 121
QY 212 VTFEHEYRDVWVFDHPSYHRPGSLYGNFCAGFQDNQFRYTLICAAACEAPLILELGG 271
Db 122 VTFEHEYRDVWVFDHPSYHRPGSLYGNFCAGFQDNQFRYTLICAAACEAPLILELGG 181
QY 272 YIYGONCMFYVNDWHASLVPLLAARYPYGVYRDSRSTLVIHNAHQGVPEASTYPDLG 331
Db 182 YIYGONCMFYVNDWHASLVPLLAARYPYGVYRDSRSTLVIHNAHQGVPEASTYPDLG 241
QY 332 LPPEWYGALEWPEWARRHALDKGEAVNFKAGVTVADRIYTVSOGYSWEVTTABGGOG 391
Db 242 LPPEWYGALEWPEWARRHALDKGEAVNFKAGVTVADRIYTVSOGYSWEVTTABGGOG 301
QY 392 LNELLSRKSVLNGIYVNGIDINDWNPATDKCPCHYSVDDLSGKAKCAELKELGLPVR 451
Db 302 LNELLSRKSVLNGIYVNGIDINDWNPATDKCPCHYSVDDLSGKAKCAELKELGLPVR 361
QY 452 EDVPLIGFGRLDYQKIDILKMAPELMREDVQFVMLGSGDPIFEGWMRSTESSYKDKF 511
Db 362 PDVPLIGFGRLDYQKIDILKMAPELMREDVQFVMLGSGDPIFEGWMRSTESSYKDKF 421
QY 512 RGVWGSFVPSVSHRITAGCDILLMPSEFPCGLNQLYAMOYGVVPHVHTGGLRDVTFEN 571
Db 422 RGVWGSFVPSVSHRITAGCDILLMPSEFPCGLNQLYAMOYGVVPHVHTGGLRDVTFEN 481
QY 572 PEGAKEEGTGNFASPLTVDKM 593
Db 482 PEGAKEEGTGNFASPLTVDKM 503

RESULT 4
US-60-312-544-9109
Sequence 9109, Application US/60312544
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkle, Gregory J.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua

TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-10(52726)A
CURRENT APPLICATION NUMBER: US/60/312,544
CURRENT FILING DATE: 2001-08-15
NUMBER OF SEQ ID NOS: 10730
SEQ ID NO 9109
LENGTH: 435
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3150-035-F10_FLI
US-60-312-544-9109

Query Match 54.1%; Score 2189; DB 24; Length 435;
Best Local Similarity 91.3%; Pred. No. 1.4e-181;
Matches 397; Conservative 16; Mismatches 22; Indels 0; Gaps 0;
QY 159 DVCGLPILAAARGHRVWVMPRYLNGSSDKNYAKALYTAKHKIPCFGGSHEVTFHEHY 218
Db 1 DVCGLPILAAARGHRVWVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHEHY 60
QY 219 RONVDKVFVDHPSYHRPGSLYGNFCAGFQDNQFRYTLICAAACEAPLILELGGYIYGNC 278
Db 61 RDSVDKVFVDHPSYHRPGSLYGNFCAGFQDNQFRYTLICAAACEAPLILELGGYIYGNC 120
QY 279 MFVVDWHASLVPLLAARYPYGVYRDSRSTLVIHNAHQGVPEASTYPDGLPPEWYG 338
Db 121 MFVVDWHASLVPLLAARYPYGVYRDSRSTLVIHNAHQGVPEASTYPDGLPPEWYG 180
QY 339 ALEWFPPEWARRHALDKGEAVNFKAGVTVADRIYTVSOGYSWEVTTABGGOGNELLS 398
Db 181 ALEWFPPEWARRHALDKGEAVNFKAGVTVADRIYTVSOGYSWEVTTABGGOGNELLS 240
QY 399 RKSVLNGIYVNGIDINDWNPATDKCPCHYSVDDLSGKAKCAELKELGLPVRDVLIG 458
Db 241 RKSVLNGIYVNGIDINDWNPATDKCPCHYSVDDLSGKAKCAELKELGLPVRDVLIG 300
QY 459 FGRLDYQKIDILKMAPELMREDVQFVMLGSGDPIFEGWMRSTESSYKDKFGRWVGS 518
Db 301 FGRLDYQKIDILKMAPELMREDVQFVMLGSGDPIFEGWMRSTESSYKDKFGRWVGS 360
QY 519 VPSVSHRITAGCDILLMPSEFPCGLNQLYAMOYGVVPHVHTGGLRDVTFENFPGAKE 578
Db 361 VPSVSHRITAGCDILLMPSEFPCGLNQLYAMOYGVVPHVHTGGLRDVTFENFPGAKE 420
QY 579 EGTGNFASPLTVDKM 593
Db 421 EGTGNFASPLTVDKM 435

RESULT 5
US-09-606-304-10
Sequence 10, Application US/09606304
GENERAL INFORMATION:
APPLICANT: Kossmann, Jens
Springer, Franziska
Abel, Gernot

TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/606,304
; FILING DATE: 28-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,567
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-606-304-10

```

```

Query Match 49.3%; Score 1992; DB 20; Length 641;
Best Local Similarity 67.5%; Pred. No. 3.9e-164;
Matches 360; Conservative 80; Mismatches 87; Indels 6; Gaps 4;

QY 102 LLEGIADS--IDSIIVAAEQD--SEIMDANEQPAQKTRSTVFVTGEAAPYAKSGGL 157
DB 90 LIPSHVAGDATWESHDIANDRODLSDETEEMEETPIKLTNFIETAAEAPYSKTGGL 149

QY 158 GDVCSGLPITALAARHVRVWVMPYKNGS--SDKNYAKALYAKHIKIPCFGSHVEYFFH 216
DB 150 GDVCSGLPMAAARHVRVWVSPRYLNGGSPDEKYANAVDLDVRATVHCFGDAQEVAFYH 209

QY 217 EYRDNDVWVFDHPSYHRPGSLYGNFCAGDNQFRYTLCCYAAEAPLILELGGYIYQ 276
DB 210 EYRAGDWFVVDHSSYCRGTPYGDYIYAGFNGQFRFTLLSHAACEAPLVLPGLGFIYGE 269

QY 277 NCMFYNDWHAASLVPLVLAARYPYGVYRDSRSTLVTHNLAHQGVPEASTYPDGLGLPEW 336
DB 270 KCLFLANDWHAALVPLVLAARYPYGVYKARSIVAIHNIHQGVPEAVTYNNLGLPPQW 329

QY 337 YGALWVPEWARHIALDKGEAVNFKGAVVTADRIYTVSOGYSWEVTTAEGGOGNELL 396
DB 330 YGAVEMIFPTWARHALDTGETVNVKGAIAVADRILTYSOGYSWEITTPEGGYGLHELL 389

QY 397 SSRKSVLNGIVNGIDINDWNPTTKCLPHYSVDDLSGKAKCABELOKELGLPVREDVPL 456
DB 390 SSRQSVLNGITNGIDVNDWNPSTDEHIAHSHYINDLSGKVOCKTDLOKELGLPIRDCPL 449

QY 457 IGFIKRLDYQKIDITKMAIPELMREDVOFVMLGSGDPIFEGWMKSTSSSKDFRGWG 516
DB 450 IGFIKRLDYQKIDITKMAIPELMREDVOFVMLGSGEKQYEDMDRHTENLFDKDFRWWG 509

QY 517 FSVPSVSHRTACGDTLLMPSRPEPCGLNOLYAMGYTPVPHVHGTGLRDTVETRNPFCAK 576
DB 510 FNVPSVSHRTACGDTLLMPSRPEPCGLNOLYAMRYGTIPVHSTGGLRDTVKDFNPYQAE 569

QY 577 G-EEGTGAFSPLTVDKMLWALRTAMSTFREHKPSWEGLMKRGMTKDHTWHA 628
DB 570 GIGEGTGWTFSPLTSEKLLDTLKLAIKGTITTEHKSSWEGLMRGMGRDYSWENA 622

```

```

RESULT 6
US-09-606-304-4
; Sequence 4, Application US/09606304
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens

```

```

; Springer, Franziska
; Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/606,304
; FILING DATE: 28-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,567
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-606-304-4

```

```

Query Match 42.5%; Score 1718.5; DB 20; Length 459;
Best Local Similarity 70.8%; Pred. No. 1.8e-140;
Matches 303; Conservative 62; Mismatches 61; Indels 3; Gaps 2;

QY 203 IPCFGGSHVEYTFREYRDNDVWVFDHPSYHRPGSLYGNFCAGDNQFRYTLCCYAAE 262
DB 14 VHCFGDAQEVAFYHEYRAGVWVVDHSSYRPGTPYGDYIYAGFNGQFRFTLLSHAACE 73

QY 263 APLILELGGYIYQNCMFVNDWHAASLVPLVLAARYPYGVYRDSRSTLVTHNLAHQGV 322
DB 74 APLVPLGGTYGKCLFLANDCNALVPLVLAARYPYGVYKARSIVAIHNIHQGV 133

QY 323 PASTYPDGLGLPEWVHGALEWVPEWARHIALDKGEAVNFKGAVVTADRIYTVSOGYSWE 382
DB 134 PAVTYNNLGLPQQYGAWEVIFPTWARHALDTGETVNVKGAIAVADRILTYSOGYSWE 193

QY 383 VTTAEGGOGNELLSRKSRLNGIVNGIDINDWNPTTKCLPHYSVDDLS--GKAKCKA 440
DB 194 ITTPEGGYGLHELLSSRQSVLNGITNGIDVNDWNPSTDEHIAHSHYINDLSRPPGKVOCKT 253

QY 441 ELQKELGLPVREDVPLTGFIKRLDYQKIDITKMAIPELMREDVOFVMLGSGDPIFEGWM 500
DB 254 DLQKELGLPIRDCPLIGFIKRLDYQKIDITKMAIPELMREDVOFVMLGSGEKQYEDMD 313

QY 501 RSTESSYKDKFRGWGVSVPVSHRTACGDTLLMPSRPEPCGLNOLYAMGYTPVPHVHGT 560
DB 314 RHTENLFDKDFRWWGVNVPVSHRTACGDTLLMPSRPEPCGLNOLYAMRYGTIPVHST 373

QY 561 GGLRDTVETRNPFCAKGE--EGTGWAFSPLTVDKMLWALRTAMSTFREHKPSWEGLMKRG 619

```



```
Db 175 NLAFIANDHWTALLPVYLKAYYRDHGLMKYTRSVLVIHNTAHOGCRPIDFRYTDLPYH 234
QY 337 YGALEWVPEWARRHALDKGEAVNFKLGAVVTADRIVTSQGSWEVTTAEGQGLNELL 396
Db 235 IDLFKLYDPV-----GGEHENIESAGLKAADRIIVTVSHGVAWEIKTSEGCGWGLHGII 286
QY 397 SSRKSVLNGVINGIDINDWNPTTKCLPH-----HYSVDDL-SGKAKCKAELQELGLPVR 451
Db 287 NENDMKLGRVINGIDITKWNPKIDVHLKSDGYNTYLETLSQSKRQCKAALQELGLPVR 346
QY 452 EDVPLIGFTGRLOYQKIDILKMAIPELMREDVQFVMLGSGDPIFGGWRMSTESSYKDKF 511
Db 347 EDVPLIGFTGRLOYQKIDILKMAIPELMREDVQFVMLGSGDPIFGGWRMSTESSYKDKF 406
QY 512 RGWVGSVPVSHRITAGCDILMPSRFEPCGLNQLYAMQYGVTVVHGTTGGLRDTVETFN 571
Db 407 RGWVGSVPVSHRITAGCDILMPSRFEPCGLNQLYAMQYGVTVVHGTTGGLRDTVETFN 466
QY 572 PFCAKEEGTGWAFSPLTVDKMLMALRTAMSTFREHKPSWEGLMKRGMTKDHWDHA 628
Db 467 PF---BESGLGWTFSATETNKLINALGNCLLTPTROYKQSWEGLRQRRGMTQDLSWDNA 520

RESULT 10
US-09-402-254-51
; Sequence 51, Application US/09402254
; GENERAL INFORMATION:
; APPLICANT: Keeling, Ganping
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; FILE REFERENCE: 2461-52
; CURRENT APPLICATION NUMBER: US/09/402,254
; EARLIER FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: PCT/US98/06660
; EARLIER FILING DATE: 1998-04-03
; EARLIER APPLICATION NUMBER: 60/042,939
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Zea mays
US-09-402-254-51

Query Match 30.3%; Score 1225.5; DB 18; Length 670;
Best Local Similarity 41.2%; Pred. No. 3.5e-97; Mismatches 97; Indels 81; Gaps 16;
Matches 277; Conservative

QY 12 APSVRLRADPATAAASACVVRARLRRLAR-GRY---VAELSREGPAARPAQQQLAPPL 67
Db 10 APPER-SGDAARLPARRNNAVSKRRDPLQVGRYSGATGTATGAAS--CQNAALADVE 66

QY 68 VPQFLAPPPAPAPQSPAP-TQPLPDAGVGEALAPDLL-----EGIAED 110
Db 67 IKSIVAAPTSTIVKFPAPGYRMLPS---GDIAPETVLPAPKPLHESPAVDGSDNGIAPP 123

QY 111 SIDSIIIVASEQDSEIMDANEQOAK-----VTRSIIVF 143
Db 124 TVEPLVQEAATWDPKYYIGDEPEADKDSRVGADDAAGSFEHYGNDGSGPLAGENVNVIV 183

QY 144 VTGEAAPAYAKSGGLGVDCSLPTALAAAGHRVNVMPVRLNGSSDKNYAKALYATAKHKI 203
Db 184 VRAECSFWCKTGGLDVGVGALPALARRGHVNVVPRY-----GDYVEAFDMGIRKYY 237

QY 204 PCFGGSHVTFPHEYRDNDVWVFDVHPSY-HRPGSLYGNFGAFGNQFRYTLCLYAAACE 262
Db 238 KAAGQDLEVNYHFAFTDGVDFVFDAPLFRHRODDIYG---GSRQETMKRMILFCRAVE 294

QY 263 APLILELGVYIGO-NCMFVNDWHASLVPLVLAARYPVGVYRDSRSTLVHNLAHQGV 321
```

```
Db 295 VPHVPCGGVCYGGNLFVFIANDHWTALLPVYLKAYYRDHGLMQYTRSVLVIHNTAHQGR 354
QY 322 EPASTYPTDLGLPEWYGALEWVPEWARRHALDKGEAVNFKLGAVVTADRIVTSQGSY 381
Db 355 GPVDEPFYMDLPEHYLQHFLYDPV-----GGEHANIFAAGLKMADRVVTVSRGYLW 406
QY 382 EVTTAEGQGLNELLSSRKSVLNGVINGIDINDWNPTTKCLPH-----HYSVDDL-SGKA 436
Db 407 ELKTVGEGWGLHDIIRSNWKNKINGIVNGIDHQBENPNKVDVHLKSDGYNTYLETLDAGKR 466
QY 437 KCAELQKELGLPVREDVPLIGFTGRLOYQKIDILKMAIPELMREDVQFVMLGSGDPIF 496
Db 467 QCAALQRELGLEVRDDVPLGLIGRUDGQGVDIIGDAMPWTAGQDVQLVMLGTGRADL 526
QY 497 EGMWSTESSYKDKFRGWGVSFVPSHRTITAGCDILMPSRFEPCGLNQLYAMQYGVTV 556
Db 527 ERMQLHLERHPNKRGVGWGVSFVPMARHRTITAGADVLVMPSRFEPGGLNQLYAMAYGVTV 586
QY 557 VHTGGGLRDTVETFNPFCAKEEGTGWAFSPLTVDKMLMALRTAMSTFREHKPSWEGLMK 616
Db 587 VHAVGGLRDTVAPDFPF---GDAGLGWTFDRAEANKLIEALRHCLDITYRKYGESWKSLOA 643
QY 617 RGMTKDHWTWDA 628
Db 644 RGMQDLSDWDA 655

RESULT 11
US-09-606-304-8
; Sequence 8, Application US/09606304
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/606,304
; FILING DATE: 28-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,567
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
```

US-09-606-304-8

Query Match
Best Local Similarity 30.0%; Score 1215; DB 20; Length 767;
Matches 264; Conservative 78; Mismatches 188; Indels 48; Gaps 12;

QY 79 PAQSPAPTOPPL-----PDAGVGEAPDILLEGIA-----EDSIDSIIAAS 120
DB 195 PPEPKSQTLLDVSNSKSLVDVPGKKIOSYPSLKESSASHIVEORNELENSSAEAN 254
QY 121 EODSEIMDNEPOAKV---TRSIIVFTGEAAPYAKSGGLDVGCSGLPIALAAARGHRVMV 177
DB 255 EETEDPVNIDEKPPPLAGTNMNIILVASECAPWSKGTGLDVGALPKALARRGRVMV 314
QY 178 VMPRYLNGSSDKYAKALYAKHILKIPCFGSGSHEVTFEYHRDNRDVFVD-HPSYHRPG 236
DB 315 VAPRYDNPQDSSG-----VRKIYKVD--GODVEVTFQAFIDGVDFVFDISHMFRIGN 368
QY 237 SLXGDNFGAGDNQFRYTLACYACEAPLILELGGYIYGQ-NCMFVNDWHASLVVLLA 295
DB 369 NIYGN---RVDILKRMVLFCKAAIEVPHVPCGGVYCGDNLVFIANDHHTALLPYLK 425
QY 296 AKYRPYGVYRDSRSTLVIHNLHOGVEPASTYPDLGLPPEWYGALEWVFPFWARRHALDX 355
DB 426 AYTRDNGIMYTRSVLVIHNLHOGGRPLEDFSYVDLPPHYMDPFKLYDPV-----G 477
QY 356 GEAVNFKGAVTADRVITVSGYSWEVTTAEGGOLNELLSSRKSVLNGIVNGIDNDW 415
DB 478 GEHFNFAAGLKTADRVITVSHGYSWELKTSQGGWGLHQIINENDWKLQGIIVNGIDTKEW 537
QY 416 NPTTKCLPH-----HYSVDDL-SGKAKCAELQKELGLPVREDVPLIGTIGRLDYQKID 470
DB 538 NPELDVLQSDGYMNYSLDTLOTGKPOCKAALQKELGLPVREDVPLIGTIGRLDPQKVD 597
QY 471 LIKMAIPELMAREDOVFMGLSGDPIFEGWMRSTESSYKDKFERGWGFSVPVSHRITAGCD 530
DB 598 LIAEASAMWQDVLVMLGTGRDLQEMURQECQHNDKIRGWGFSVKTSHRITAGAD 657
QY 531 ILLMPSRFEPCGLNQLYAMQYTPVVVHGTGGLRDTVETNPFEGAKGEETGNWAFSPLTV 590
DB 658 ILLMPSRFEPCGLNQLYAMQYTPVVVHGTGGLRDTVETNPFEGAKGEETGNWAFSPLTV 590
QY 591 DKMLWALRTAMSTFREHKPSWGLMKRGMTKDHWDHA 628
DB 715 SOLIHALGNCLLYRETKKSWEGIQTRCMTQDLSWDNA 752

RESULT 12
US-09-402-254-49
; Sequence 49, Application US/09402254
; GENERAL INFORMATION:
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; FILE REFERENCE: 2461-52
; CURRENT APPLICATION NUMBER: US/09/402,254
; EARLIER FILING DATE: 1999-10-01
; EARLIER FILING DATE: 1998-04-03
; EARLIER FILING DATE: 1998-04-03
; EARLIER FILING DATE: 1997-04-04
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 49
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Zea mays
US-09-402-254-49

Query Match 29.9%; Score 1209.5; DB 18; Length 804;

Best Local Similarity 41.9%; Pred. No. 1.2e-95;
Matches 288; Conservative 85; Mismatches 228; Indels 87; Gaps 18;

QY 12 APSVRLRADPA-----TAARASACVVRARLRRLARGRYVAELSRREGPAARPAQOQOLAPP 66
DB 56 AASVRAAAAAGSESEPAKSSSS-SQAGAVQGSTAKAVDSAPPNPLTSAPKOSASAM 114
QY 67 L--VPGEFLAPPPAPAPQSPAPTQPLP-----DAG-----VGLAPDLILLGIAEDSI 112
DB 115 QNGTSGSSASTAAPVSGPKADHPSAPVTKREIDASAVKPEPAGDDARPVESIGIAE-PV 173
QY 113 DSIIVASEODS-----EIMDANEQPOA-KVTRSIIVFTGEAAPYAKSGGLDVGCSGLP 165
DB 174 DAKADRAPATDAASAPYDREDNEPGLAGPNVNMVVVAVASECAPKCTKGTGLDVGVALP 233
QY 166 IALAARGHRVVMVMPRYLNGSSDKYAKALYAKHILKIPCFGSGSHEVTFEYHRDNRDVMV 225
DB 234 KALARRGHRVMVPIRY-----GEYAEARDLGVRRRYKVAGQDSEVTFYFHSIDGVDFV 287
QY 226 FVDHPSY-HRPGLSYLGDNFGAGDNQFRYTLACYACEAPLILELGGYIYGQ-NCMFVNVN 283
DB 288 FVEAPPFRHNNIYG---GERLDILKRMILFCKAAVEVPYAPCGGTVYVYDGNLVEIAN 344
QY 284 DHASIVPVLAAKYPYGVYRDSRSTLVIHNLHOGVEPASTYPDLGLPPEWYGALEWV 343
DB 345 DHTALLPVYLKAYYRONGLMQYARSVLVIHNLHOGRPVDFVDFVDFVDFVDFVDFVDFV 394
QY 344 FPEWARRHALDK-----GEAVNFKGAVTADRVITVSGYSWEVTTAEGGOLNELL 395
DB 395 -----EHYIDHFPLKLDNIGGDHSHNVFAAGLKTADRVITVSNYMWELKTSSEGGWGLHDI 448
QY 396 LSSRKSVLNGIVNGIDNDWNTTKCLPH-----HYSVDDL-SGKAKCAELQKELGLPV 450
DB 449 INDNWKLQGIIVNGIDNDWNTTKCLPH-----HYSVDDL-SGKAKCAELQKELGLPV 508
QY 451 REDVPLIGTIGRLDYQKIDLIKMAIPELMAREDOVFMGLSGDPIFEGWMRSTESSYKDK 510
DB 509 RDDVPLIGTIGRLDYQKIDLIKMAIPELMAREDOVFMGLSGDPIFEGWMRSTESSYKDK 568
QY 511 FRGWGFSVPVSHRITAGCDIILLMPSRFEPCGLNQLYAMQYTPVVVHGTGGLRDTVET 570
DB 569 VRAWGFSVPLAHRTAGADILLMPSRFEPCGLNQLYAMQYTPVVVHGTGGLRDTVET 628
QY 571 NPFAGKGEETGNWAFSPLTVDKMLWALRTAMSTFREHKPSWGLMKRGMTKDHWDHAP- 629
DB 629 DPF---NDTGLGWTDFDRAENRMDALSCHLTYRNYKESWRACRARGMAEDLSWDHAAV 685
QY 630 -----SSTRSSSGSPSWT 642
DB 686 LYEDVLVKAKYQWANLATRRRSCRTWT 713

RESULT 13
US-09-625-406-11
; Sequence 11, Application US/09625406
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/625,406
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/941,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 698 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-625-406-11

Query Match 29.8%; Score 1205.5; DB 20; Length 698;
Best Local Similarity 43.0%; Pred. No. 2.1e-95;
Matches 283; Conservative 83; Mismatches 221; Indels 71; Gaps 17;
QY 12 APSVRLRADPA-----TAAASACVVRARLRRLARGRYVAELSREGPAARPAQOQOLAPP 66
DB 56 AASVRAAAAPAGSESEAAKSSSS--SQAAGVQSGTAKAVDSAPPNPLTAPKQSQSAAAM 114
QY 67 L--VPGFLAPPPAPAPASPTAPPLP-----DAG-----VGLAPDLLLLLEGIADSI 112
DB 115 QNGTSGSSASTAAPVSGPADHSPVTKREIDASAVKPEPADGADARPVESIGIAE--PV 173
QY 113 DSIIVASEQDS-----ETMDANEOPQA--KVTFESIYFVTGEAAPYAKSGGLGDCVCSLP 165
DB 174 DAKADAAPATAAASAPYREDNEPGLAPGNVNMVVVAVASECAPCKTGLGDDVVGALP 233
QY 166 IALAARGHRVNMVPRYLNGSSDKNYAKALYAKHIKIPCGGSHSEYTFHEYRDNDYDW 225
DB 234 KALARRGHRVMVPIPR-----GEYAEARDLGVRRYKRVAGQDSEVTYFHSYIDGVDF 287
QY 226 FVDHPSY--HRPGSLYGNFAGFQNGQRYTLACYAAEAPLLELGGYIYGO--NCMEVFN 283
DB 288 FVEAPPFRHRHNTYG---GERLDLKRMLLFCKAAVEVPMYAPCGGTIVYGDGNLVEFAN 344
QY 284 DWHASLPVLLAAKRYPGVYRDSRSTLVINLHAGQVEPASTYPDGLGPPWYGALEWV 343
DB 345 DWHALLPVLYLKAYRDNLQYARSVLVINIAHQGRGPVDDFVDFDLP----- 394
QY 344 FPEWARRHALDK-----GEAVNFKGAVVTADRVITVSQGYSWETVTAEGGQGLNEL 395
DB 395 -----EHYIDHFKLYDNIGDHSNVFAAGLKTADRVVTVSNGYMWELKTSEGGWGLHDI 448
QY 396 LSSRKSVLNGIVNGIDINDWNPPTDKCLPH-----HYSVDDL--SKAKCKAELOKELGLPV 450
DB 449 INQNDWKLOGLVNGIDSEWNPADVHLHSDDYNTYTFETLDTGRKQCKAALQOLGLQV 508
QY 451 REDVPLIGFGRLDYQKIDILKMAIPELMREDVQFVLMGSGDPIFEGWNRSTESSYKDK 510
DB 509 RDDVPLIGFGRLDHQKGVDIADAIHWIAGODVQLVMLGTGRADLEDMLRFESEHSKD 568
QY 511 PRGWGFSVPVSHRITAGCDILLMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVET 570
DB 569 VRWYGFSPVLAHRTAGADILLMPSRPEPCGLNQLYAMAYGTVPVHAVGGLRDTVAPF 628
QY 571 NPEKAKGEGGTGAFSPVLTVDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHTWDHA 628
DB 629 DPF-----NDTGLWTFDRAEARMIDALSHCLITTYNYKESNRACRARGMAEDLSWDHA 683

RESULT 14
US-09-388-743-26

; Sequence 26, Application US/09388743
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Typha latifolia
; US-09-388-743-26

Query Match 29.6%; Score 1196.5; DB 17; Length 801;
Best Local Similarity 42.6%; Pred. No. 1.6e-94;
Matches 269; Conservative 80; Mismatches 200; Indels 83; Gaps 15;
QY 42 GRVVAELSRGPAARPAQOQOLAP-----PLVPGFLAPPPP-----APAQSPAPT 86
DB 109 GKYS-----GAVPNYSQLAODDTSENPLVNSFGGSPKDNVEAVEFQVRSVADAFG 161
QY 87 QPPLPDAGVAGELAPDLLLEG-----TAEDSIDSIIVAASEQDSEIMDANEQP- 133
DB 162 RPEPSLGTTKILSPFYLEAESDCAENAEDELVEAKLDSHVH---KDDLNPGEENEVPL 218
QY 134 --QAKVTRISVFTGEAAPYAKSGGLGDCVCSLPALAAARGHRVNMVPRYLNGSSDKNY 191
DB 219 PLAGANYMNIIVAAECAPMSKTKGGDGVAGALPKALARGHRVNMVPRYGNTPAEQDI 278
QY 192 AKALYAKHIKIPCGGSHSEYTFHEYRDNDYDWVFDHPSY--HRPGSLYGNDFCAFQDQ 250
DB 279 GVRKYKVVH-----GQDMEVTFYHAYIDGVDFVMDSPDFRHRGNRIYEGN---RVDIL 329
QY 251 FRYTLACYAAEAPLLELGGYIYGO--NCMFVNDWHASLPVLLAAKRYPGVYRDSRS 309
DB 330 KRMILFCKAAVEVPMVHPVPCGFCYGDGNLAFITNDNHTALLPVLYLKAYYRDNLMLKYARS 389
QY 310 TLVTHNLHAGQVEPASTYPDGLGPPWYGALEWVPEWARRHALD-----KGEAVNF 361
DB 390 VLVTHNLHAGQVEPDDFKFVGLPD-----HYLDLFLYDPVCGEHLNI 433
QY 362 LKGAVVTDRIYTVSQGYSWETVTAEGGQGLNELLSRKSVLNGIVNGIDINDWNPPTDK 421
DB 434 FAAGLKTADRVITVSHGYAWELKTSEGGWGLHEINESNNKFKQGIYNGIDAKEWSPEDV 493
QY 422 CLPH-----HYSVDDL--SKAKCKAELOKELGLPVREDVPLIGFGRLDYQKIDILKMAI 476
DB 494 HLKSDGYTNSLDPTLENGKPVCKAALQREVGLPVDRDNPVPIAFIIGRLDHQKGVDLIAEAM 553
QY 477 PELMREDVQFVLMGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPS 536
DB 554 PWIVSHDVQVVMGLGTGRQDLENLLRNEFGQHRDKAVRAVAESVKMAHRTAGADILLMPS 613
QY 537 RFEPCLNQLYAMQYGTVPVHGTGGLRDTVETPNFGAKGEEGTGNAFSPVLTVDKMLWA 596
DB 614 RFEPCLNQLYAMMYGTIPVHVAVGGLRDTVTFQDPF---NESGLGWTDFDRAEAKGLIHA 670
QY 597 LRTAMSTFREHKPSWEGMLKRGMTKDHTWDHA 628
DB 671 LNNCLNTYWNYSKSNKGLQTRGMMQDLSWDNA 702

RESULT 15
US-09-606-304-6
; Sequence 6, Application US/09606304
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; Springer, Franziska

Abel, Gernot
TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/606,304
FILING DATE: 28-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,567
FILING DATE: <Unknown>
APPLICATION NUMBER: DE P 44 41 408.0
FILING DATE: 10-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: Agrevo-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 558 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-606-304-6

```

Query Match          29.3%; Score 1184; DB 20; Length 558;
Best Local Similarity 46.3%; Pred No. 1.le-93;
Matches 257; Conservative 76; Mismatches 186; Indels 36; Gaps 11;

QY 90 LPDAGVCELAPDLLEGTA-----EDSIDIIVAASQDSEIMDANEPOQAKV---TRS 140
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
9 VPGKKIOSYMPSLRKESSASHVEQRNENLEGSAAENEETDPYNIDEKPPPLACTNVMN 68
QY 141 IVFTGTGAAPYAKSGGGLDVCGSLPIALAAARGHRVVMVMPRYLNGSSDKNYAKALYAKH 200
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 IILVASCAPWKTGGGLDVGAGLPKALARAGHRVVMVAPRYDNYPEPDGSG----VRKI 124
QY 201 IKIPCGGSHVETFFHEYRDNDVWVFVD-HPSYHRPGSLYGDNFGAFGDNQFRYTLACYA 259
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 YKVD--QGDVEVYTFQAFIDGVDFVIDSHMFRHIGNNIYGGN---RVDILKRMVLFEKA 179
QY 260 ACEAPLILELGGYIYGO-NCMFVVDNWHASLPVLLAAKYRPGYVYDRSSTLIHNLH 318
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 AIEVPWHVPGGVCYGDGNLVFIANDWHTALLPVLLKAYYRDNGIMNYTRSLVVIHNAH 239
QY 319 QGVPEASTYDPLGLPEWTYGALESWFPEWARRHALDKGEAVNFKGAVVTADRVIVTSQG 378
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 QGRGPLEDSYVDLPHYMDPFKLYDPV-----GGEHFNIFAAGLCTADRVAVTVSHG 291
QY 379 YSWEVTTAGGQGLNELLSSRKSVLNGIVNGIDINDWNPETDKCLPH-----HYSVDL-S 433
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
292 YSWEKLTISQGGWGLHGIINENDWKLOGIVNGIDTKEMNPELVHLQSDGYMNYSLDTLQT 351
QY 434 GKAKCAELQKELGPLRVEDYPLIGFIRLDYQKGDIGDILKMAIPELWREDYQVFWLGGSD 493
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

11

us-09-674-824-2.rapm

Fri Mar 29 09:26:42 2002

OM of: US-09-674-824-2 to: GenEmbl.* out_format : pfs
Date: Mar 28, 2002 5:38 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=slp
-Q/cgn2_1/USPTO.spool/US095674824/runat_27032002_094251_18421/app_query.fasta_1.824
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOPCU=0.000 -LOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTPMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09674824_CGNL_1_6134
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIME=120 -WARN_TIME=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-674-824-2
Query length: 756
Database: GenEmbl.*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 2539.800000

score_list:

Sequence	Strd Orig	Zscore	EScore	Len	Documentation
gb_pat:AX010492	+ 4044.00	3630.23	6.3e-194	2771	AX010492 Sequence 1 from Patent
gb_pl:TAE292522	+ 3998.00	3589.88	1.1e-191	2421	AJ292522 Triticum aestivum mRNA
gb_pl:TAE292521	+ 3933.00	3531.11	2.1e-188	2575	AJ292521 Triticum aestivum mRNA
gb_pl:AF091803	+ 3930.00	3528.38	3.0e-188	2591	AF091803 Triticum aestivum mRNA
gb_pat:AX031272	+ 3930.00	3248.19	3.0e-188	2662	AX031272 Sequence 11 from Patent
gb_pat:TAU48227	+ 3612.00	3323.97	2.1e-172	2239	U48227 Triticum aestivum mRNA
gb_pl:AF15890	+ 3313.00	2976.20	1.7e-157	2055	AF15890 Oryza sativa subsp. indica
gb_pl:AF158904	+ 2829.50	2540.63	3.1e-133	2585	E06904 cDNA encoding soluble s
gb_pl:RICSS	+ 2821.50	2533.60	7.6e-133	2533	D16202 Rice mRNA for soluble s
gb_pat:AF03359	+ 2809.50	2523.25	2.9e-132	2383	AJ3359 Sequence 1 from Patent
gb_pl:AF036891	+ 2776.50	2492.05	1.6e-130	2591	AF036891 Zea mays starch synth
gb_pl:AF168786	+ 2770.50	2487.66	2.7e-130	2592	AF168786 Sorghum bicolor synth
gb_pat:AF06494	+ 2713.50	2439.23	1.4e-127	1752	AF106494 Sequence 12 from Patent
gb_pat:AR049919	+ 2651.00	2379.41	2.9e-124	2992	AJ63308 Sequence 1 from Patent
gb_pl:AF234163	+ 2632.50	2346.64	2.0e-122	8387	AF234163 Hordeum vulgare starch
gb_pat:AX031274	+ 2546.00	2099.70	2.9e-119	1620	AX031274 Sequence 20 from Patent
gb_pat:AX031274	+ 2349.00	2099.70	1.1e-108	10336	AX031274 Sequence 13 from Patent
gb_pat:AF091802	+ 2315.00	2068.51	6.1e-107	11387	AF091802 Aegilops tauschii st
gb_pat:AR112888	+ 2001.00	1797.65	7.5e-92	2360	AJ1207 Sequence 9 from Patent
gb_pl:STASANT	+ 2001.00	1797.65	7.5e-92	2360	AR112888 Sequence 9 from Patent
gb_pat:AB026295	+ 1786.00	1574.85	1.9e-79	8900	Y10416 S. tuberosum mRNA for sc
gb_pl:RICE1	+ 1756.00	1568.94	4.1e-79	8900	D38221 Rice gene for soluble s
gb_pat:AF1201	+ 1731.50	1557.81	1.7e-78	1758	AJ1201 Sequence 3 from Patent
gb_pat:AR112885	+ 1731.50	1557.81	1.7e-78	1758	AR112885 Sequence 3 from Patent
gb_pl:AF121673	+ 1526.00	1366.63	7.6e-68	4617	AF121673 Arabidopsis thaliana
gb_pl:AB006701	+ 1517.50	1338.77	2.7e-66	8194	AB006701 Arabidopsis thaliana
gb_pl:AF068834	+ 1250.50	1126.38	1.8e-54	2480	AF019297 Zea mays starch synth
gb_pl:AF068834	+ 1246.00	1120.81	3.7e-54	2097	AF068834 Ipomoea batatas starch
gb_pl:AF016493	+ 1239.50	1112.81	1.0e-53	2865	AF019296 Zea mays starch synth
gb_pl:AF395537	+ 1215.00	1090.99	1.7e-52	2793	AF395537 Oryza sativa soluble
gb_pat:AF1205	+ 1215.00	1090.99	1.7e-52	2793	AJ1205 Sequence 7 from Patent
gb_pat:AR112887	+ 1212.50	1088.14	2.5e-52	3049	AF12887 Sequence 7 from Patent
gb_pl:AF026421	+ 1212.50	1088.14	2.5e-52	3049	AF026421 Chlamydomonas reinhar
gb_pl:TAE269503	+ 1204.00	1080.65	6.5e-52	2988	AJ269503 Triticum aestivum mRNA
gb_pl:PSSTASYNT	+ 1204.00	1080.65	6.5e-52	2988	X88790 P.sativum mRNA for star
gb_pl:TAE269504	+ 1203.50	1080.67	6.4e-52	2793	AJ269504 Triticum aestivum mRNA

gb_pl:AF155217	+ 1200.50	1077.86	9.2e-52	2842	! AF155217 Triticum aestivum
gb_pl:AF383878	+ 1195.50	1073.22	1.7e-51	2903	! AF383878 Oryza sativa solub
gb_pl:AF173900	+ 1187.00	1066.43	4.0e-51	2575	! AF173900 Manihot esculenta
gb_pat:AF51203	+ 1184.00	1065.76	4.4e-51	1926	! AF51203 Sequence 5 from Pat
gb_pat:AR112886	+ 1184.00	1065.76	4.4e-51	1926	! AR112886 Sequence 5 from pa
seq_name: gb_pat:AX010492					
seq_documentation_block:					
LOCUS	AX010492	2771 bp	DNA	PAT	06-SEP-2000
DEFINITION	Sequence 1 from Patent WO9958688.				
ACCESSION	AX010492				
VERSION	AX010492.1 GI:9997335				
KEYWORDS	bread wheat.				
SOURCE	Triticum aestivum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
REFERENCE	Luetticke,S., Block,M. and Loerz,H.				
AUTHORS	Nucleic acid molecules which code for enzymes derived from wheat				
TITLE	and which are involved in the synthesis of starch				
JOURNAL	Patent: WO 9958688-A 1 18-NOV-1999;				
FEATURES	LUTETICKE STEPHANIE (DE); BLOCK MARTINA (DE); LOERZ HORST (DE); HOECHST SCHERING AGREVO GMBH (DE)				
Location/Qualifiers					
1..2771					
/organism="Triticum aestivum"					
/db_xref="taxon:4565"					
280..2550					
/note="unnamed protein product"					
/codon_start=1					
/protein_id="CAC07516.1"					
/db_xref="GI:9997336"					
/translation="MAATGCGAGCLAPSRLRADPATAARASACVVRARLRRLRGARY					
VAELSRGPAARPAQOOLAPLYPFLAPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP					
LLEGIREDSDSIITVAASEQSEIMDANEQPAQVTRSIIVFTGEEAAYAKSGGLGDL					
CGSLPITALAARGHVRVMPYRLNGSSDKNYAKALYTHAKIKIPCFGSEHVTFFHEY					
RDNDVDFVHPHSPYHRGSLYDNGFAGDNQRYTLICYAACAPLILGLGGYIYGQ					
NCMYGALWVFPWARRHALDKGEAVNFKGAVTADRTVTVSGSYNEVTAEGGGL					
NELLSRRKSVLNGIDIDNWNPTDKCLPHYVDDLSGKAKCAELQKELGLPV					
REDVPLIGTIGRLDYKIDILKAIPELMREDVQVHLGSGDIFEGWNRSTESSTYK					
DKFRGWGFSVPVSHRTIACGDIILLMSREFPCNLQIYAGQVTVPVHVTGGLRDT					
VETNPFKAGEEGTGNFSLTVYDKMLWALRTAMSTFRHKPFSWELMKRGMTKQHT					
WDHAPSSTSRSSGSPSTNCRGLGRKSPSALKTSSSFRGEGYPTLRCP					
ATVESQACALLWFAGSRTPYDCAAAVATASGRQLQFWGIRKCAAGWLTKHSDGS					
LSVRVTAEIRNLVTL"					
BASE COUNT	669 a	699 c	754 g	649 t	
ORIGIN					
alignment_scores:					
		Quality: 4044.00	Length: 756		
		Ratio: 5.349	Gaps: 0		
		Percent Similarity: 100.000	Percent Identity: 100.000		
alignment_block:					
US-09-674-824-2 x AX010492					
Align seg 1/1 to: AX010492 from: 1 to: 2771					
1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17					
280 ATFGCGGCGACGCGGCTCGGCGGCGGCTCGGCGGCGGCTCGGCGGCT 329					
17 uArgAlaAspProAlaThrAlaAlaAlaArgAlaSerAlaCysValValArgA 34					
330 GCGCGCGGATCGCGGACGCGGCGGCGGCTCGGCGGCTCGGCGGCT 379					
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50					

380 CGCGGCTCCGGCGCTTGGCGGGCGCCCTACGTCGCGGAGCTCAGCAGG 429
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
430 GAGGCT 479
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
480 CGTCCGAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
530 CCCCAGCAGCGCCCGCTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 579
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
580 GACCTCCTCGCTCGAGGAGTTCGTCGAGGATTCCATCGACAGCATAAATGT 629
117 IAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
630 GGCTGCAAGTGAGCAGGATTCGAGATCATGATCGGATCGGATCGCAACCTC 679
134 IAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
680 AAGCTAAAGTTACAGTAGCATGCTGTTGTGACTGGTGAAGCTGCTCCT 729
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
730 TATGCAAAAGTCAGGGGGTGGAGATGTTGTGCTGCTGCTACCAATATGC 779
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
780 TCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
830 ATGGGTCTCTCATAAAACTATGCAAGGCATTTATACACTGCGAAGCAC 879
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
880 ATTAAGATTCCATGCTTTGGGGATCATCATGAAGTGACCTTTTTCATGA 929
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
930 GTATAGAGACAACGTCGATTGGTGTGTTGTCGATCATCGTCATATCACA 979
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
980 GACCAGGAAGTTTATATGAGATAAATTTGGTGTCTTTGGTGATATATCAG 1029
251 PheArgTyrThrLeuLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
1030 TTCAGATACACACTCTTTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
1080 TGAATTGGGAGGATATATTTATGACAGAAATTCATGTTTGTGTGAACG 1129
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
1130 ATGGGATGCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
1180 TACGGTGTTCACAGAGATTCGCCAGCACCCCTTGTATATACATAAATTTAGC 1229
317 aHisGlnGlyValIleProAlaSerThrTyrProAspLeuGlyLeuProP 334
1230 ACATCAGGGTGTGGAGCCTCAAGTACATATCCTGATCTGGGATTCCTC 1279
334 roGluTyrTyrGlyAlaLeuGluTyrValPheProGluTyrPheAlaArgArg 350
1280 CTGAATGCTATGGAGCTTTAGATGGGTATTTCCAGAAATGGGCAAGGAGG 1329

351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
1330 CATGCCCTTGACAAGGGTCAGGAGTTAACTTTTGAAGGAGCAGTTGT 1379
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
1380 GACAGCAGATCGGATTTGACCTCAGTCAGGCTTATTCATGGGAGGTCA 1429
384 hrThrAlaGluGlyGlnGlyLeuAsnGlnLeuLeuSerSerArgLys 400
1430 CAACCTGCTGAAGGTGACAGGGCCTCAATGAGCTCTTAAGCTCCCGAAA 1479
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
1480 AGTGTATTGAATGCAATTTGTAATGCAATTAATGATTTGGAACCC 1529
417 oThrThrAspLysCysLeuProHisIleThrSerValAspAspLeuSerG 434
1530 CACCACAGACAAGTGTCTCCCTCATCATTTATCTGTCGATGACCTCTCTG 1579
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1580 GAAAGCCCAATGTAAAGCTGAATTCAGAGGAGTTGGGTTTACCTGTA 1629
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1630 AGGAGGATGCTCTCTGATTTGGCTTTTATTTGGAAGACTGGAATTTACCAGAA 1679
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgLysAspV 484
1680 AGGCATTGATCTCATTAATAATGCCCATTCAGAGCTCATGAGGAGGAGCG 1729
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500
1730 TGCAATTTGTCATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATG 1779
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
1780 AGATCTACCGAGTCGAGTTACAAGGATAAATTCGTCGATGGGTGGATT 1829
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1830 TAGTGTCCAGTTTCCACAGATAAATTCGAGGTTGCGATATATTGTTAA 1879
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1880 TGCCATCGAGATTTGAACCTTGGCTTAAATCAGCTATATGCTATGCAA 1929
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
1930 TATGTCAGATTCTCTGTAGTTTCAAGGAGGAGGAGGAGGAGGAGGAGG 2029
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
1980 CGAGACCTTCAACCTTTTGGTGCAGGAGGAGGAGGAGGAGGAGGAGG 2079
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
2030 CGTTCTACCCCTAACCGTGGACAGAGTGTGGGCAATTCGGAACCCCG 2079
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
2080 ATGTCGACATTCAGGAGCAGACAGCCGCTCTGGGAGGGGCTCATGAAGCG 2129
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSera 634
2130 AGGCATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2179
634 rgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650
2180 GATCTTCGAGTGGGCTTCGTCGGACCAACCCCTACGTCATGTAGAGGGGA 2229

```

651 LeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSe 667
|||||
2230 CTGGGAGGTCACAGTGCAGTCTCTACGCTCTGAAGACATCCTCTTC 2279
667 rSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 684
|||||
2280 ATCCTTCGCGCGCGCGGAAGATACCCCTGTACATTCGTTGCTCTGCTA 2329
684 hrValGluSerGlnCysAlaCysLeuLeuTyrPheAlaGlySerArgThr 700
|||||
2330 CAGTAGAGTCGCAATGCGCTTGTCTTGTGCTGCGCGGTTCGAGAACA 2379
701 TyrAspGlyCysAlaAlaAlaValThrAlaSerGlyArgGlnLe 717
|||||
2380 TATGACGGCTGCTGCTGCGCGGTGACAGCTTCGGGTGACACAGTT 2429
717 uGlnPheTrpGlyLeuArgLysGlyCysAlaAlaGlyTrpLeuThrAla 734
|||||
2430 ACAGTTTGGGGAATAAGGAAGGATGCTGTCAGAGTGGTTACAGCAA 2479
734 yshHisSerAspGlySerLeuSerValArgValThrAlaGluIleArg 750
|||||
2480 AGCACCACTCAGATGCGAGCTCTCTGCTGCTTACAGCTGAATCAGA 2529
751 AsnGlnLeuValThrLeu 756
|||||
2530 AACCAACTGGTGACTCTT 2547

```

seq_name: gb_pl:TAE292522

```

seq_documentation_block: 2421 bp mRNA PLN 21-JUL-2000
LOCUS TAE292522 Triticum aestivum mRNA for starch synthase I-2 (wSSI-2 gene).
DEFINITION Triticum aestivum mRNA for starch synthase I-2 (wSSI-2 gene).
ACCESSION AJ292522
VERSION AJ292522.1 GI:9369335
KEYWORDS starch synthase I-2; wSSI-2 gene.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
Peng, M., Hucl, P. and Chibbar, R.N.
1 (bases 1 to 2421)
Isolation, characterization and expression analysis of starch
synthase I from wheat (Triticum aestivum L.)
Unpublished
2 (bases 1 to 2421)
Chibbar, R.N.
Direct Submission
Submitted (18-JUL-2000) Chibbar R.N., Cereal Biotechnology, Plant
Biotechnology Instt., Natl. Research Council Canada, 110 Gymnasium
Place, Saskatoon, Saskatchewan, S7N 0W9, CANADA
Location/Qualifiers
1..2421
/organism="Triticum aestivum"
/cultivar="Felder"
/db_xref="taxon:4565"
/dev_stage="12-14 days-post-anthesis"
/tissue_type="developing kernels"
127..2295
/gene="wSSI-2"
127..2070
/gene="wSSI-2"
/EC_number="2.4.1.21"
/function="essential for starch synthesis"
/codon_start=1
/product="starch synthase I-2"
/protein_id="CAB99210.1"
/db_xref="GI:9369336"
/translotion="MAATGAGCCLAPSVRLRADPATAARASACVVRRLRRVARGRY
VAELSRREGAPAAQQOOLAPLPVPGFAPPAQSPAPTOPPLPDAGVGELAPDL
LLEGIAEDSIDSIIVAAEQDSEIMDAKQOAKVTRSIIVFTGEARAPAKSGGLGDV
CGSLPIALAAHRHVRVMPRYLNGSSDKYKALYAKHIKIPCFGGSEVTFFFHY

```

FEATURES

source

gene

CDS

```

RDNDVDFVDFHPSYHRPGSLYGDNEGAFGDNOFRVTLTYCAACEAPLILELGGYLYGO
NCMFVNDWHSALVPVLLAARYPYGVYRDSRSTLV IHNLAHOGVEPASTYPDLCGLPP
EMYGALVWFPEWARHRLDKGEAVNFKLGAVVADRIVTVYVSOQYSWEVTTAEGQGL
NELLSRSKVLNGIYNGIDINDWNPTDKCLPHYSVDDLSGKAKCAKAELOKLEGLPV
REDVPLIFIGRLDYOKGIDLIKMAIPELMRDVOFVLMGSDPIFEFGMRSTESSYK
DKFRGWFVSPVSHRITAGDILLMPRSFRPCGGLNQLYAMQYGVTVVYVHGTGLRDT
VETFNPFKAGEEGTGAWSPLTYVDKMLWALRTAMSTFREHKPSWEGLMKRGMTKDHT
WDHAAEQYEQIFEWAFVQPYVM"
transit_peptide 127..250
/mat_peptide
/gene="wSSI-2"
251..2067
/gene="wSSI-2"
/product="starch synthase I-2"
2290
/gene="wSSI-2"
2290..2295
/gene="wSSI-2"
BASE COUNT 539 a 602 c 708 g 572 t
ORIGIN
alignment_scores:
Quality: 3998.00 Length: 757
Ratio: 5.302 Gaps: 1
Percent Similarity: 99.604 Percent Identity: 99.207
alignment_block:
US-09-674-824-2 x TAE292522
Align seg 1/1 to: TAE292522 from: 1 to: 2421
1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
|||||
127 ATGGCGCGGACGCGGCTCGCGCGGCTCGCTCGCGCGGCTCGCGCT 176
17 uArgAlaAspProAlaThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 34
|||||
177 GCGCGCGGATCGCGCGGCGCGGCGCGGCGCGGCTCGCTCGCTCGCGG 226
34 laArgLeuArgArgLeuAlaAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
227 CGCGGCTCGCGGCGGCTCGCGGCGGCGGCTCGCTCGCGGCTCGCAGG 276
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 67
|||||
277 GAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 326
67 uValProGlyPheLeuAlaProProProProProAlaProAlaGlnSerPro 84
|||||
327 CGTGCCAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 376
84 laProThrGlnProProProProProProProProProProProProPro 100
|||||
377 CCCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
101 AspLeuLeuLeuGluGlyTyrAlaGluAspSerIleAspSerIleIleVa 117
|||||
427 GACCTCTCTGCTCGAAGGATTCCTGAGGATTCCTGAGGATTCCTGAGG 476
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
|||||
477 GCGTCAAGTGAGCAGGATTCGTGAGATCATGATGATGATGATGATGATG 526
134 InAlaAlaValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
527 AAGCTAAAGTTACTCGCAGCATCGCTGTTGTTGTTGTTGTTGTTGTTG 576
151 TyrAlaLysSerGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
577 TATGCAAGTACGCGGGGTGGGAGATGTTGTTGTTGTTGTTGTTGTTG 626
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
|||||

```

627 TCTTGCTGCTCGTGTCCACCGAGTGATGGTTGTAATGCCAAGATACCTTAA 676
184 sNGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
677 ATGGGCTCTCTGATTAATAAACTATGCAAGGATATATACACTGGGAACAC 726
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217
727 ATTAAAGATTCCATGCTTTGGGGATCATCATGAAGTGAACCTTTTTCATGA 776
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
777 GTATAGACAACTGATGGGTTTGTTCGATCATCCGTCATATCACA 826
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
827 GACCAGGAAGTTTATATGGAGATAATTTGGTGCTTTTGGTGATTAACAG 876
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
877 TTCAGATACACACTCCCTTTGCTATGCTGCATGCCAGGCCCACTAATCCT 926
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
927 TGAATTGGGAGGATATATTTATGGACAGATTCATGTGTTTGTGAACG 976
284 spTTPHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300
977 ATTGCATGCCAGCCTTGTGCCAGTCCCTTCCTGCTGCAAAATATAGACCA 1026
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
1027 TAGCGTGTACAGAGATCCCGCAGCACCTTGTGTATACATAATTTAGC 1076
317 ahISGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334
1077 ACATCAGGGTGTGGAGCCTGCAAGTACATATCCTGATCGGGATTGCCCT 1126
334 toGluTyrTyrGlyAlaLeuGluTyrPValPheProGluTyrAlaArgArg 350
1127 CTGAATGGTATGGAGCTTTAGAAATGGGTATTTCCAGATGGCAAGGAGG 1176
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
1177 CATGCCCTTGACAAGGGTGAGCAGTTAACTTTTGAAGGAGCAGTGTGT 1226
367 ThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
1227 GACAGCATCGGATGTGACCGTCAGTCAGGGTTATTCTATGGGAGGTCA 1276
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1277 CAACGTGTAAGGTGGACAGGCCCTCAATAGCTCTTAAGCTCCCGAAAA 1326
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
1327 AGTGATTTGAATGGAAATGTAATGGAAATGGAAATGGAAATGGAAATGG 1376
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspLeuSerG 434
1377 CACCACAGACAGTGTCTCCCTCATCATATTTCTGTCGATGACCTCTCTG 1426
434 lYlYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1427 GAAAGGCCAAATGTAAGCTGAATTTGCAGAGGAGTTGGGTTTACCTGTA 1476
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1477 AGGGAGGATGTTCCTCTGATTTGCTTTTATTTGAAGACTGGATTACAGAA 1526
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspY 484
1527 AGGCATTGATCTCATTAATAATGGCCATTCAGAGCTCATGAGGAGGAGC 1576

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
1577 TGCAATTTGTTCATGCTGGATCTGGGGATCCCAATTTTGAAGGCTGGATC 1626
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
1627 AGATCTACCGAGTCGAGTTTACAAGGATAATCCCGTGGATGGGTTGGATT 1676
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1677 TAGTGTTCAGATTTCCACAGATAACTGCAGGTTGGATATATTTGTTAA 1726
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1727 TGCCATCGAGATTTGAACCTTGGCTTAAATCAGCTATATGCTATGCAA 1776
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
1777 TATGGTACAGTTCCTGTAGTTTCATGGAACCTGGGGGCTCCGAGACACAGT 1826
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
1827 CGAGACCTTCAACCTTTTGGTGCANAAAGGAGAGGAGGTACAGGTTGG 1876
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1877 CGTTCTCAGCGCTAACCGTGGACAAAGATGTTGTGGCATTTGCCAACC 1926
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1927 ATGTCGACATTCAGGAGACACAGCCGCTCTGGGAGGGGCTCATGAACGG 1976
617 qGlyMetThrLysAspHisThrTrpAspHisAla . ProSerSerThrSer 633
1977 AGGATGACGAAAGACCATACGTGGGACCATGCCGCCGAGCAGTACGAGC 2026
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
2027 AGATCTTCGAGTGGGCTTCTGTCGACCAACCCCTACGTCATGTAGACGGG 2076
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
2077 ACTGGGAGGTCCCAAGTCCGAGTCTCTCTTGAGCTCTGAAGACATCCTCT 2126
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683
2127 CATCTTCCGCGGCCGGAAGGATACCCCTGTACATTTGCCGTTGCTCTGCT 2176
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700
2177 ACAGTAGAGTCGCAATGCGCCTGCTTGTCTTGGTTCCGCCGTTCCAGAAC 2226
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717
2227 ATATGACGGCTGTGCTGCTGCGCGGTGCACAGCTTCGGGTGGAGCAGCT 2276
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733
2277 TACAGTTTGGGGGATAAGGAAGGATGTGCTGAGGATGGTTAACAGCA 2326
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750
2327 AAGCACCCTCAGATGGCAGCTCTCTGCTCGTGTACAGCTGAATCAG 2376
750 GasnGlnLeuValThrLeu 756
2377 AAACCAACTGGTGACTCTT 2395

seq_name: gb_pl:TAE292521

seq_documentation_block:

LOCUS TAE292521 2575 bp mRNA PLN 21-JUL-2000

DEFINITION

Triticum aestivum mRNA for starch synthase I-1 (wSSI-1 gene).

ACCESSION

AJ292521

GI:9369333

KEYWORDS

starch synthase I-1; wSSI-1 gene.

SOURCE

bread wheat.

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.

REFERENCE

AUTHORS

Peng, M., Hucl, P. and Chibbar, R.N.

TITLE

Isolation, characterization and expression analysis of starch synthase I from wheat (Triticum aestivum L.)

JOURNAL

UNPUBLISHED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .2575

/organism="Triticum aestivum"

/cultivar="Fielder"

/db_xref="taxon:4565"

/dev_stage="12-14 days-post-anthesis"

/tissue_type="developing kernels"

223. .2290

/gene="wSSI-1"

223. .2166

/gene="wSSI-1"

/EC_number="2.4.1.21"

/function="essential for starch synthesis"

/codon_start=1

/product="starch synthase I-1"

/protein_id="CAB99209.1"

/db_xref="GI:9369334"

/translation="MAATGVAGCLAPSRLRADPATAARASACVVRRLRLARGRY

VAELSRGPAARPAQQQLAPFLVFCFLAPPPAPAPQAPPTQPPDPAGVGLAPDL

LLGAEADSIDIIIVAAQSEIDANDKQAKVTRISIVFTGEAAYAKSGGLGV

CSSLPTALAAHRVVMVMPRYLNGSDKNYAKLYTGKHKIPCGSHVFTFHEY

RQNVDFVVDHPSYHRPGSLYDNGFCADQNFRTLLCYACEAPLILELGGYLYGO

NCMFVNDWHAISLVPVLLAAKYRPYGVYRDSRTLVIIHLAQGVPEASTYDPLGPP

ENYGALEWFFPEARRHALDKGEAVNFKGAVVTADRIVTVSQGYSWERTVAEGQGL

NELLSRKSVLNGIVNGIDINDNPTDKLPHYSVDLDSGAKCKAELOKELGPV

REDVPLTIGRLDYQKGIIDLKMAIPELMREDYQVFMVLSGDPIFEQWMRSTESSYK

DFRQWVGFSVPVSHRTAGCDLILMPSRFPCGLNQLYAMQYGVVYVHGTGLRDT

VETFPNFGAKGEGTGWAFSPILTVDKMLWALRTANSTFREHKPSHEGLMKRGMTKDHT

WDHAAEOYEOIFEMAFVDQPYVM"

transit_peptide 223. .346

/gene="wSSI-1"

/evidence=experimental

347. .2163

/gene="wSSI-1"

/evidence=experimental

/product="starch synthase I-1"

2285. .2290

/gene="wSSI-1"

BASE COUNT 567 a 654 c 748 g 606 t

ORIGIN

alignment_scores:

Quality: 3933.00

Ratio: 5.251

Percent Similarity: 98.943

Percent Identity: 98.547

Length: 757

Gaps: 1

alignment_block:

US-09-674-824-2 x TAE292521

Align seg 1/1 to: TAE292521 from: 1 to: 2575

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17

```

|||||
223 ATGGCGGACGAGGGCGTGGCGCGGGTGCCTCGCCCCACGCGTCCGCT 272
|||||
17 uArgAlaAspProAlaThrAlaAlaAtgAlaSerAlaCysValValArgA 34
|||||
273 GCGGCGGATCCGGAGCGCGCGCGCGCGTCCGCTCGCTCGCGCG 322
|||||
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
323 CGCGGCTCCGCGCGTGGCGGGCGCGTACGTCGCGAGTCAGCAGG 372
|||||
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProLe 67
|||||
373 GAGGCGCGCGCGCGCGCGCGCGCAGCAGCAGCACTGGCGCGCGCT 422
|||||
67 uValProGlyPheLeuAlaProProAlaProAlaGlnSerProA 84
|||||
423 CGTCCAGGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 472
|||||
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
473 CCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 522
|||||
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleLeVa 117
|||||
523 GACCTCTGCTCGAAGGGATGCTGAGGATTCATCGACAGCATAATTGT 572
|||||
117 laAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
|||||
573 GGCTGCAAGTGAGCAGGATTCGAGATCGGATGCGAATGAGCAACCTC 622
|||||
134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
623 AAGCTAAGTTACACGTCAGTCTGTTGTGACTGTTGTAAGCTGCTCCT 672
|||||
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
673 TATGCAAGTCAGGGGGGCTGGGAGATGTTGGTTCGTACCAATTGC 722
|||||
167 aLeuAlaAlaArgGlyIleArgValMetValValMetProArgTyrLeuA 184
|||||
723 TCTTGTGCTGCTGGTGCACCGTGTGATGTTGTAATGCCAAGATACTGA 772
|||||
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
|||||
773 ATGGGTCCTCTGATAAAACTATGCAAGGCGATTTATACACTGGGAAGCAC 822
|||||
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
|||||
823 ATTAAGATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGA 872
|||||
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
|||||
873 GTATAGAGAACAGTCGATTGGGTGTTTTCGATCATCGTCATATCATPA 922
|||||
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||
923 GACCAGGAAGTTTATGAGATAATTTTGGTGCTTTGGTGATAATCAG 972
|||||
251 PheArgTyrThrLeuLeuCystTyrAlaAlaCysGluAlaProLeuIleLe 267
|||||
973 TTCAGATACACACTCTTTTGTCTATGCTGCATCGAGGCGGCCACTAATCT 1022
|||||
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
|||||
1023 TCAATTGGGAGGATATATTATGGAGCAATTCATGTTGTTGTTGAACG 1072
|||||
284 sPTpHisAlaSerLeuValProValLeuAlaAlaLysTyrArgPro 300
|||||
1073 ATTGCATGCCAGCCTTGTGCGAGTCCCTTCTGCTGCAAAATATAGACCA 1122
|||||
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||

```

1123	TACGGTGTTTACAGAGATTCCCGCAGCACCCCTTGTATACATATTTAGC	1172	617	gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer	633
317	aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp	334	2073	AGCATGACGAAGACCATACGTGGGACCATGCCGCCGAGCAGTACGAGC	2122
1173	ACATCAGGGTGTGAGCGCTGCAAGTACATATCTCTGATCTGGGATTGCCAC	1222	634	ArgSerSerSerGlyProSerThrTrpAsnProThrSerCysArgArgG1	650
334	roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg	350	2123	AGATCTTCGAATGGCGCTTCGTGGACCAACCCCTACGTCTAGACGGGG	2172
1223	CTGAATGGTATGGAGCTTTAGAAATGGGTATTTCCAGAAATGGGCAAGGAGG	1272	650	yLeuGlyArgSerLysCysGlnSerProSerAlaLeuLysThrSerSerS	667
351	HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa	367	2173	ACTGGGAGGTTCAGCGGGTCTCTTCAGCTCTGAAGACATGTCCT	2222
1273	CATGCCCTTGACAGGGGTGAGCGAGTTAACTTTTGAAGGAGCAGTCGT	1322	667	erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla	683
367	lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT	384	2223	CATCTTCGCCGCCCGGAAGGATACCCCTGTACATTGCGTTGTCCTGCT	2272
1323	GACAGCAGATCGAATTTGTGACCGGTGACGTCAGGGTTATTTCATGGGAGGTCA	1372	684	ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh	700
384	hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys	400	2273	ACAGTAGAGTCGCAATGCGCTGCTTGCT.TGGTCCGCCGTTTCGAGAGT	2321
1373	CNACTGCTGAAGGTGGACAGCGCCCTCATGAGCTCTTAAGCTCCGAAAA	1422	700	rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL	717
401	SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr	417	2322	AGATGACGGCTGCTGCTCGCGGGTGACAGCTTCGGGTGGATGACAGT	2371
1423	AGTGATTGAATGGAATTTGTAATGGAATTGACATTAATGATTGGAACCC	1472	717	euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla	733
417	oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG	434	2372	TACAGTTTGGGAATAAGGAAGGATGTGCTGCAGGATGGTTAAACAGCA	2421
1473	CACCACAGCAAGTGTCTCCCTCATCATATTATCTGCGATGACCTCTCTG	1522	734	LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr	750
434	lYlsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal	450	2422	AAGCACCTCAGATGCGACCTCTCTGTCGGTGTACAGCTGAAATCAG	2471
1523	GAAGGCCAAATGTAAGCTGAATTTGCAGAGGAGCTGGGTTTACCTGTA	1572	750	gAsnGlnLeuValThrLeu	756
451	ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy	467	2472	AAACCAACTGGTGACTCTT	2490
1573	AGGGAGGATGTTCCCTGATTGGCTTTATTTGGAAGACTGGATTACCAAA	1622	seq_name:	gb_pl:AF091803	
467	sGlyLeuAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV	484	seq_documentation_block:	2591 bp	mrna
1623	AGGCATTGATCTCATTAATAATGGCCATCCAGAGCTCATGAGGAGGAGC	1672	LOCUS	AF091803	PLN 14-SEP-1999
484	aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet	500	DEFINITION	Triticum aestivum starch synthase I mRNA, complete cds.	
1673	TGCAGTTTGTCTGCTGCTGATCTGGGATCCCAATTTTGAAGCTGGATG	1722	ACCESSION	AF091803	
501	ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh	517	VERSION	AF091803.1	GI:5880465
1723	AGATCTACCGAGTCGACTTACAAAGGATAATTCCTGGATGGTTGGATT	1772	KEYWORDS	bread wheat.	
517	eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM	534	SOURCE	Triticum aestivum	
1773	TAGTGTTCAGTTTCCACAGATAACTCAGGTTGCGATATATTGTTAA	1822	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
534	etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln	550	REFERENCE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
1823	TGCCATCCAGGTTGAACCTTTGGTCTTAATCAGCTATATGCTATGCAA	1872	AUTHORS	Poideae; Triticeae; Triticum.	
551	TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa	567	TITLE	1 (bases 1 to 2591)	
1873	TATGGTACAGTCTCTGTAGTTTCAAGTACGAGGCTCCGAGACACAGT	1922	JOURNAL	Cloning and characterization of a gene encoding wheat starch	
567	lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA	584	THEOR. Appl. Genet. 98, 1208-1216 (1999)		
1923	CGAGACCTTCAACCTTTTGGTCAAAAGAGAGAGGAGGTACAGGTTGGG	1972	2 (bases 1 to 2591)		
584	laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla	600	AUTHORS	Li, Z., Rahman, S., Kosar-Hashemi, B., Mouille, G., Appels, R. and	
1973	CGTCTCACCCTTAACCGTCGACAGATGTTGTGGCATTTGCAACCGCG	2022	Morell, M.K.		
601	MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr	617	Direct Submission		
2023	ATGTCGACATTCAGGGAGCACAAGCGCTCTCGGGAGGGGCTCATGAAGCG	2072	Submitted (10-SEP-1998)		
			1600, Canberra, ACT 2601, Australia		
			Location/Qualifiers		
			1. .2591		
			/organism="Triticum aestivum"		
			/db_xref="taxon:4565"		
			/chromosome="7"		
			/tissue_type="endosperm"		
			256. .2199		
			/codon_start=1		
			/product="starch synthase I"		
			/protein_id="AAC54661.1"		
			/db_xref="GI:5880466"		
			CDS		
			source		
			1. .2591		
			/organism="Triticum aestivum"		
			/db_xref="taxon:4565"		
			/chromosome="7"		
			/tissue_type="endosperm"		
			256. .2199		
			/codon_start=1		
			/product="starch synthase I"		
			/protein_id="AAC54661.1"		
			/db_xref="GI:5880466"		

```

/translation="MAATGVGAGCLAPSRLRADPATAARASACVVRARLRRLARGRY
VAELSRGCPAARPAQOOLAPPLYPGLAPPPAPQSPAPTQPLPDAGVGLAPDL
LLEGIATEDSIDIIAASEQDSESSDNEQPOAKVTRSIIVFTGEAAYAKSGGLGDV
CGSLPIALAAAGHRVMVMPRYLNGSSDKNYAKALYTKHKIKIPCFGSHSVTFPHY
RDNVDMFVDPSPHVRPGSLYGDNFGAFDQNRFTLLCYAAACEAPLILELGGYIYGO
NCMEVNDWHSALVPVLLAAKYRPGYVYRDSRSLVTHNLAHOGLEPASTYDGLGLPP
EMYGALWEFPFEMARRHLDKGEAVNFKLGAVTADRIVTVSOGYSWEVTTASGGGGL
NELSSRSKSLVNLGVINGIDINDMNPTDKCLPHYSVDDLSGRKAKKAELOKELGLPV
REDVPLIGFTIGRLDYOKGIDLIKWAPELMRDQVQFVLMGGDPIFEQWMRSTESSYK
DKRGWVGFSVPVSHRITACGDLILMFSEPCGLNOLYAMQYGTVPVVGHTGGLRDT
VETENPFCAGGEBGTGWAFLSPLTVDKMLMALRTAMSTFREHKPSWEGLMKRGMTKDHT
WDHAAEQIQLFENAFVDQPYVM"
BASE COUNT      567 a      672 c      743 g      609 t
ORIGIN

```

```

alignment_scores:
  Quality: 3930.00      Length: 757
  Ratio: 5.247          Gaps: 1
  Percent Similarity: 98.943      Percent Identity: 98.415

alignment_block:
US-09-674-824-2 x AF091803      ...
Align seg 1/1 to: AF091803 from: 1 to: 2591

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
|||||
256 ATGGCGGGGACGGCGCTGGCGGGCGGTCCTCGCCCCCGCGGCTCGCGCT 305

17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
306 GCGCGCCGATCGCGGCGCGCGCGCGCGCGCTCGCGCTCGCGCTCGCGG 355

34 laArgLeuArgArgLeuAlaArgGlyArgGlyValAlaGluLeuSerArg 50
|||||
356 CGCGGCTCGCGGCTGGCGCGGGCGCGCTACGTTGCGGAGCTCAGCAGG 405

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
|||||
406 GAGGCGCGCGCGCGCGCGCGCGCGCGCAGCAGCAGCAACTGGCGCGCG 455

67 uValProGlyPheLeuAlaProProProAlaProAlaProAlaGlnSerProA 84
|||||
456 CGTGCCAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 505

84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
506 CCGCGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 555

101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
|||||
556 GACCTCTGCTCGAAGGATGCTGAGGATTCATCGACAGCATATATGT 605

117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
|||||
606 GGCTGCAAGTGACAGGATTCAGATCATGGATGCGAATGACCAACCTC 655

134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
656 AAGCTAAGTTACAGTAGCATGCTGTTGTGACTGGTGAAGCTGCTCT 705

151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
706 TATGCAAGTCAAGGGGGCTGGGAGATGTTGTGGTTCGTTACCAATGCG 755

167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
|||||
756 TCTTGCTGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 805

184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
|||||
806 ATGGGTCTCTGATATAAACTATGCAAGGCATTATACACTGGGAGCAGC 855

```

```

201 lleLysIleProCysPheGlySerHisGluValThrPhePheHisG1 217
|||||
856 ATTAAGATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGA 905

217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
|||||
906 GTATAGACACAACCTCGATTGGGTGTTTTCGATCATCCGTCATATCAT 955

234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||
956 GACCAAGAAAGTTATATGGAGATAATTTTGGTGTCTTTGGTGAATCAG 1005

251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
|||||
1006 TTCAGATACACACTCCCTTTCCTATGCTGCATGCGAGGCCCTTAATCCT 1055

267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
|||||
1056 TGAATTGGGAGGATATATTTATGACAGAATTGCATGTTTGTGTGAACG 1105

284 sPTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
|||||
1106 ATTGGCATGCCAGCCTTGGCCAGTCTCTTCTGCAAAATATAGACCA 1155

301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||
1156 TAGCGTGTTTTACAGAGATTCCCGCAGCACCTTTGTTTATACATAATTTAGC 1205

317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334
|||||
1206 ACATCAGGCTCTGGAGCCTGCAAGTACATATCTGATCTGGGATTGCCAC 1255

334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
|||||
1256 CTGAATGGTATGAGCTTTAGAAATGGGTATTTCCAGATGGCGCAGGAGG 1305

351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||
1306 CATGCCCTTTGACAAGGGTGAGGAGTAACTTTTGAAGGAGCAGTCTGT 1355

367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
|||||
1356 GACAGCAGATCGAATTGTGACCGTCAGTCAGGCTTATTCATGGGAGGTCA 1405

384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||
1406 CAACCTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAA 1455

401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||
1456 AGTGATTGATGGAAATTTGAATGGAATTGACATTAAATGATGGAAACCC 1505

417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
|||||
1506 CACCACAGACAAGTGTCTCCCTCATCATATTCTGTGATGACCTCTCTG 1555

434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
|||||
1556 GAAAGGCCAAATGTAAGCTGAATTCAGAAAGGAGCTGGGTTTACCTGTA 1605

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
|||||
1606 AGGAGGAGTGTCTCTGATTGGCTTTTATTTGGAAGACTGGATTACCAAG 1655

467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||
1656 AGGCATTGATCTCAATAAATGGCCATTTCCAGAGCTCATGAGGAGGAGC 1705

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500
|||||
1706 TGCAGTTTGTCTGCTGGATCTGGGGATCCCAATTTTGAAGGCTGGATG 1755

```

Fri Mar 29 09:26:43 2002

```
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGly 517
|||||
1756 AGATCTACCGAGTCGAGTACAAGGATAAAATCCGTGGATGGGTGGATT 1805
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
|||||
1806 TAGTGTTCACAGTTCCACACAGAACTACGAGTGGATATATTGTTAA 1855
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMeGln 550
|||||
1856 TGCCATCCAGGTTGAACCTTGTGCTTAATCAGCTATATGCTATGCAA 1905
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||
1906 TATGTGACAGTTCCTGTGTATGATGGAAGTGGGGCCCTCCGAGACAGT 1955
567 IGLuThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
|||||
1956 CGAGACCTTCAACCCCTTTTGGTGTCAAAAGGAGAGGAGGGGTACAGGTGG 2005
584 laPheSerProLeuThrValAspLysMetLeuTyrAlaLeuArgThrAla 600
|||||
2006 CGTCTCACCGCTAACCCGTGGACAAGATGTTGTGGGCATTGCGAACCCG 2055
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||
2056 ATGTGACATTCAGGAGCACAAAGCCCTCTGGGAGGGGCTCATGAAGCG 2105
617 gGlyMetThrLysAspHisThrTrpAspHisAla..ProSerSerThrSer 633
|||||
2106 AGGCATGACGAAGACATPACGTGGGACCATTGCCCGCAGCAGTACGAGC 2155
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
|||||
2156 AGATCTTCGAATGGGCTTCGTGGACCAACCTACGTATGCTAGACGGG 2205
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
|||||
2206 ACTGGGAGGTCGAAGCGGGGTCTCCTTGAGCTCTGAAGACATGTTCT 2255
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683
|||||
2256 CATCTTTCGCGCGCCGGAAGGATACCCCTGTACATTGCTGTCTCTGCT 2305
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700
|||||
2306 ACAGTAGAGTCGAATGCGGCTGCTTGTCT..TGGTCCGCGGTTGAGAGT 2354
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717
: |||||
2355 AGATGACGGCTGTGCTGCTGCGCGGTGACAGCTTCGGGTGGATGACAGT 2404
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733
|||||
2405 TACAGTTTTGGGGAATGAAGAGGATGCTGCTCAGGATGGTTAACAGCA 2454
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluLeuAr 750
|||||
2455 AAGCACCACTCAGATGGCAGCCTCTCTGCTCCGTGTTACAGCTGAATCAG 2504
750 gAsnGlnLeuValThrLeu 756
|||||
2505 AAACCAACTGGTGACTCTT 2523

seq_name: gb_pat:AX031272
seq_documentation_block:
LOCUS AX031272 2662 bp DNA PAT 20-SEP-2000
DEFINITION Sequence 11 from Patent WO9914314.
ACCESSION AX031272
VERSION AX031272.1 GI:10278604
KEYWORDS Aegilops tauschii.
SOURCE
```

ORGANISM Aegilops tauschii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Aegilops.
REFERENCE 1 (bases 1 to 2662)
AUTHORS Li,Z., Morell,M. and Rahman,S. In plants
TITLE Regulation of gene expression
JOURNAL Patent: WO 9914314-A 11 25-MAR-1999;
GOODMAN FIELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;
RAHMAN SADEOUR (AU) ; UNIV AUSTRALIAN (AU) ; COMM SCIENT IND RES
ORG (AU) ; GROUPE LIMAGRAIN PACIFIC PTY L (AU)
FEATURES
source 1..2662
/organism="Aegilops tauschii"
/db_xref="taxon:37682"
/tissue_type="ENDOSPERM"
misc_feature 1..2651
/product="NUCLEOTIDE SEQUENCE OF CDNA WHEAT SSS 1"
BASE COUNT 592 a 677 c 754 g 639 t
ORIGIN
alignment_scores:
Quality: 3930.00 Length: 757
Ratio: 5.247 Gaps: 1
Percent Similarity: 98.943 Percent Identity: 98.415
alignment_block:
US-09-674-824-2 x AX031272 ..
Align seg 1/1 to: AX031272 from: 1 to: 2662
1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
|||||
257 ATGGCGCGACGGGCGTCGGCGCGGGGTGCTCCGCCACAGCTCGCCCT 306
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
307 CGCGCGCGATCCGCGAGCGGGCGCGGCGTCCGCTGCTGCTCGCG 356
34 laArgLeuArgArgLeuAlaArgGlyTyrValAlaGluLeuSerArg 50
|||||
357 CGCGGCTCCGCGCTTGGCGGGGCGCTACGTTGCCGAGCTCAGCAGG 406
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
|||||
407 GAGGCGCGCGCGCGCGCGCGCGCGCAGCAGCACTGGCCCGCGCGCT 456
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
|||||
457 CGTGCCAGGCTTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
507 CCGCGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 556
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
|||||
557 GACCTCTCTGCTGAAGGATGCTGAGGATTCATCGACAGCAGATAATGT 606
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnPro 134
|||||
607 GGCTCAAGTGAGCAGGATTCCTGAGATCATGATCGGAATGAGCAACCTC 656
134 InAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
657 AAGCTAAAGTTACAGTACATGCTGTTGTGACTGGTGAAGCTGCTCT 706
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
707 TATGCAAAAGTCAGGGGGGCTGGGAGATGTTTGTGTTGCTTACCAATTGC 756
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184

|||||
757 TCTTGTGCTCGTGGTCAACCGTGTGATGGTTGTAATGCCAAGACTTGA 806
|||||
184 snglySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
|||||
807 ATGGGTCCTCTGATAAAACTATGCAAGGCATTATACACTGGGAAGCAC 856
|||||
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217
|||||
857 ATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTTTCATGA 906
|||||
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
|||||
907 GTATAGAGACAACGTCGATTGGGTGTTGTCGATCATCCGTCATATCATA 956
|||||
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||
957 GACCAGGAAGTTATATGGAGATAAATTTGGTGCTTTTGGTGATAATCAG 1006
|||||
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
|||||
1007 TTCAGATACACACTCCTTGTGATGTCGATGGAGGCCCCCACTAATCCT 1056
|||||
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
|||||
1057 TGAATTGGGAGGATATATTATGGACAGAATTGCATGCTTTGTTGTGAACG 1106
|||||
284 spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300
|||||
1107 ATGGGATGCCACCTTGTGCCAGTCTCTTGTGTCGCAAAATATAGACCA 1156
|||||
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||
1157 TACGGTGTTTACAGAGATTCCCGCAGACACCTTGTATACATAATTAGC 1206
|||||
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
|||||
1207 ACATCAGGGTCTGGAGCCTGCAAGTACATATCTGATCTGGGATTGCCAC 1256
|||||
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
|||||
1257 CTGAATGGTATGAGCTTTAGATGGGTATTCAGAATGGGCAAGGAG 1306
|||||
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||
1307 CATGCCCTTGACAAGGCTGAGGCAGTTAACTTTTGAAGGACAGTCGT 1356
|||||
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
|||||
1357 GACAGCAGATCGAATTTGTGACCCGTCAGTCAGGGTTATTTCATGGGAGGTCA 1406
|||||
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||
1407 CAACTGCTGAAGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAATA 1456
|||||
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||
1457 AGTGTATTGAATGGAATTTGAATGGAATGACATTAATGATTGGAACCC 1506
|||||
417 oThrThrAspLysCysLeuProHisIstYrSerValAspAspLeuSerG 434
|||||
1507 CACCACAGACAGTGTCCTCATATTCTGTCGATGACCTCTCTG 1556
|||||
434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
|||||
1557 GAAAGGCCAAAATGAAGCTGAATTCGAGAAGGAGCTGGGTTTACCTGTA 1606
|||||
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
|||||
1607 AGGGAGGATGTTCTCTGATGGCTTTATTGGAAGACTGGATTACCAGAA 1656
|||||
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||

1657 AGCATTGTATCTCATTAATAATGGCCATTCCAGAGCTCATGAGGGAGCACG 1706
|||||
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||
1707 TGCAGTTTGTTCATGCTTGGATCTGGGATCCAATTTTGAAGGCTGGATG 1756
|||||
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
|||||
1757 AGATCTACCAGAGTCGAGTTACAAGGATAAATCCGTGGATGGGTGGATT 1806
|||||
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
|||||
1807 TAGTGTTCAGTTTCCACAGAATAACTGCAGGTTGCGATATATGTTAA 1856
|||||
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||
1857 TGCATCTCCAGGTTTGAACCTTGTGGTCTTAATCAGCTATATGCTATGCAA 1906
|||||
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||
1907 TATGGTACAGTTCTCTGTAGTTTCATGGAACCTGGGGGCTCCGAGACACAGT 1956
|||||
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584
|||||
1957 CGAGACCTTCAACCTTTTGGTGCARAAAGGAGAGGAGGTACAGGTGGG 2006
|||||
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||
2007 CGTGTCTCACCGCTAACCGTGGACAAGATGTTGGGCATTTGCGAACCGCG 2056
|||||
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||
2057 ATCTCAGACATTAGGAGGACACAGCCGCTCTGGAGGGGCTCATGAGCG 2106
|||||
617 gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer 633
|||||
2107 AGGCATGACGAAAGACCATACGTGGGACCATGCCCGCAGCAGTACGAGC 2156
|||||
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
|||||
2157 AGATCTTCGAATGGGCTTCGTGGACCAACCTTACGTCATGTAGACGGG 2206
|||||
650 yLeuGlyArgSerLysCysCysLeuSerProSerAlaLeuLysThrSerSerS 667
|||||
2207 ACTGGGAGGTCGACGCGGGTCTCTTTCAGCTCTGAAGACATGTTCT 2256
|||||
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683
|||||
2257 CATCTTCCGCGCGCGGAGGATACCCCTGTACATTGCGTTGCTCTGCT 2306
|||||
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700
|||||
2307 ACAGTACAGTCGCATGCGCTCTGCT.TGGTCCGCGCGGTTTCGAGAGT 2355
|||||
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717
|||||
2356 AGATGACGCTGTGCTGCTCGCGCGGTGACAGCTTCGGGTGGATGACAGT 2405
|||||
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733
|||||
2406 TACAGTTTTGGGAATAAGGAAGGATGTGCTGCAGGATGGTTAACAGCA 2455
|||||
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750
|||||
2456 AAGCACCACTCAGATGGCAGCCTCTCTGTCCTGTTTACAGCTGAATTCAG 2505
|||||
750 gAsnGlnLeuValThrLeu 756
|||||
2506 AAACCAACTGGTGACTCTT 2524
|||||

seq_name: gb_pat:A93350

seq_documentation_block:

LOCUS A93350 2239 bp DNA 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9745545.
ACCESSION A93350
VERSION A93350.1 GI:6741617
KEYWORDS bread wheat.
SOURCE Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 2239)
Block.M. and Loerz.H.
NUCLEIC ACID MOLECULES ENCODING ENZYMES FROM WHEAT WHICH ARE
INVOLVED IN STARCH SYNTHESIS
Patent: WO 9745545-A 1 04-DEC-1997;
HOECHST SCHERING AGREVO GMBH (DE); BLOCK MARTINA (DE)
FEATURES
Location/Qualifiers
1..2239
/organism="Triticum aestivum"
/strain="CV. FLORIDA"
/db_xref="taxon:4565"
/clone="TASSS"
/haplotype="CA. 21 D CARYOPSES"
/clone_lib="CDNA LIBRARY IN PBLUESCRIPT SK (-)"
<3..2018
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB69543.1"
/db_xref="GI:6741618"
/translation="TQPPLPDAGVELAPDLLEGIAEDSIDSIIVAAEQSEIMDA
NEPOAKYTRSVFTGEAAYKAGSGLDVCGSLPIAALAAHRHVMVMPYRLNGSS
DKNYAKALYTKHKIKPCRGSGHEVTFHEYRDNDVDFVDPHSYHRPGSLYDNFGA
FDGNFRYLLCYAAACEAPLLELGGYIVGQNCMFVNDWHASLVPLLAARYPYGV
YRDSRTLVHNLHQGVPEASTYDGLPPEWYGVALEWVPEWARRHALDKGEAVNF
LKAVYADRIYVSGYSWEVTTABGGGGLNELSSRSVLNGIVNGIDINDWNPPT
LCKPHYSVDLSGKAKCKAEQLKELGPVREDVPLIGFIRLDITQKIDILKMAIP
ELMRDVPVGLSGDPIPEGMRSSTESSYKDKFRWGVFSPVSHRITAGCDILLMP
SRFPCGLNOLYAMQYGTVPVYVHTGGLRDITVETNPFKAGEEGTGWAFSPLTDKM
LWALRTFMFTKHPKPSWGLMKRMTKDHWTDPASSTSRSSSGSPSWNPNTSCRRGL
GRSKCESALATSSSRPGPEGYPCTLRCPATVESQACLLWFAFSRTYDCAAAAV
TASGGRLQFWIRKGCACAGWLTAKHHSDLSLVRTAIEINQLVTL"
BASE COUNT 611 a 448 c 590 g 590 t
ORIGIN
alignment_scores:
Quality: 3612.00 Length: 671
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-674-824-2 x A93350 ..
Align seg 1/1 to: A93350 from: 1 to: 2239
86 ThrGlnProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102
3 ACGAGCGCGCCCTGCGGACGCGCGGTGGGGAACCTCGCGCCGACCT 52
102 uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAla 119
53 CCTGCTCAAGGAGATTGCTGAGGATTCATCGACAGCATATATTGGCGTG 102
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAla 135
103 CAAGTGACGAGATTCTGAGATCATGATGCGATGACACACCTCAAGCT 152
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152
153 AAAGTTACACGTAGCATCGCTTTGTGACTGCTGGAAGCTGCTCCTTATGC 202
152 aLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuA 169
|||||

203 AAAGTCAGGGGGGTGGAGATGTTTGTGGTTCGTTACCAATTCCTCTTG 252
169 laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185
|||||
253 CTGCTCGTGGTCAACCGAGTGATGCTTGTAAATGCCAAGATACTTAAATGGG 302
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLys 202
|||||
303 TCCCTCTGATAAAACTATGCAAGGCAATTATACACTGCGAAGCACATTAA 352
202 siLeProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrA 219
|||||
353 GATTCCATGCTTTGGGGATCAGATGAAGTACCTTTTTCATGAGTATA 402
219 rgAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 235
|||||
403 GAGACAACGTCGATTGGGTGTTTTCGATCATCCGTATATACACAGACCA 452
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252
|||||
453 GGAAGTTTATATGGAGATAATTTTGGTCTCTTTGGTGATATCATGTTTCA 502
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269
|||||
503 ATACACACTCCTTTGCTATGCTGATCGAGGCGCCCTCAATCTTGAAT 552
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrp 285
|||||
553 TGGGAGGATATATTTATGGACAGAAATTCATGTTTGTGTAACGATTGG 602
286 HisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGl 302
|||||
603 CATGCCACCTTGTGCCAGTCCTTCTTCTGCAAAATATAGACCATACGG 652
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
|||||
653 TGTTTACAGAGATTCCCGCAGACACCTTGTATACATAATTTAGCACAT 702
319 lnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProGlu 335
|||||
703 AGGTGTGGAGCCCTGCAAGTACATATCCTGATCTGGGATTCCTCTGAA 752
336 TrpTyrGlyAlaLeuGluTyrValPheProGluTyrAlaArgArgHisAl 352
|||||
753 TGGTATGGAGCTTTAGAATGGGTATTTCCAGAATGGCAGAGGAGCATGC 802
352 aLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrA 369
|||||
803 CCTTGACAAGGGTGAGGAGTAACTTTTGAAGGAGCAGTTGTGACAG 852
369 laAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThr 385
|||||
853 CAGATCGGATTGTGACCGTCAGTCAGGTTTATTCATGGGAGTCACAACT 902
386 AlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402
|||||
903 GCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGT 952
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrT 419
|||||
953 ATTGAATGGAATGTAAATGGAATTTGACATTAATGATTGGAAACCCACCA 1002
419 hrAspLysCysLeuProHisIstYrSerValAspAspLeuSerGlyLys 435
|||||
1003 CAGACAAGTGTCTCCCTCATCATATTATCTGTCGATGACCTCTCTGGAAAG 1052
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGl 452
|||||
1053 GCCAAATGTAAGCTGGAATTCAGAAGGAGTTGGGTTTACCTGTAAAGGA 1102
452 uAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469
|||||
1103 GGATGTTCTCTGATGGCTTTATTTGGAAAGACTGGATTACCAAGAAAGCA 1152

```

469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485
|||||
1153 TTGATCTCATTAATGATGCCATCCAGAGCTCATGAGGGAGGACGTGCAA 1202
486 PheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgse 502
|||||
1203 TTTGTCATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATCAGATC 1252
502 rThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerV 519
|||||
1253 TACCCAGTCGAGTTACAAGGATAAATCCGTGGATGGGTTGATTTAGTG 1302
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535
|||||
1303 TTCCAGTTTCCACACAGAATACTGCAGGTCCGATATATTGTTAATGCCA 1352
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG 552
|||||
1353 TCGAGATTTGAACCTTCGGGTCTTAATCAGCTATATGCTATGCAATATGG 1402
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGlu 569
|||||
1403 TACAGTTCTCTAGTTGATGAACCTGGGGCTCCGAGACACAGTCGAGA 1452
569 hrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpAlaPhe 585
|||||
1453 CTTCAACCCCTTTGGTGCAAAAGGAGAGGAGGTACAGGTGGCGTTC 1502
586 SerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSe 602
|||||
1503 TCACCGCTAACCGTGGACAGATGTTGTGGCATTTGCAACCCGATGTC 1552
602 rThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyM 619
|||||
1553 GACATTCAGGAGACACAGCGCTCTGGAGGGCTCATGAACGGAGGCA 1602
619 etThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSer 635
|||||
1603 TGACGAAGACCATACCTGGGACCATGCCCGGACGAGTAGACAGATCT 1652
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG 652
|||||
1653 TCGAGTGGGCTTCGTGGACCAACCTACGTCATGTAGACGGGACTGGG 1702
652 YArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerP 669
|||||
1703 GAGTCCAGTCCGAGTCTCTTCAGCTCTGAAGACATCTCTTCATCCT 1752
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685
|||||
1753 TCGCGCGCGGGAAGGATACCCCTGTACATTGCGTTGCTCTACAGTA 1802
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702
|||||
1803 GAGTCCGAATGCGCCCTGCTGCTTTGGTTCGCGGTTTCGAGAATATGA 1852
702 pGlyCysAlaAlaAlaValThrAlaSerGlyArgGlnLeuGlnP 719
|||||
1853 CGGCTGCTGCTGCGCGCGTGACAGCTTCGGGTGACACAGTATACAGT 1902
719 heTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
|||||
1903 TTTGGGGAATAAGGAAGGATGTGTCAGAGTGTTAACACCAAGCAC 1952
736 HisSerAspGlySerLeuSerValArgValThrAlaGluIleArgAsnG 752
|||||
1953 CACTCAGATGGCAGCTCTCTGCTCGGTGTACAGCTGAATCAGAAACCA 2002
752 nLeuValThrLeu 756
|||||
2003 ACTGGTGACTCTT 2015

```

seq_name: gb_pl:TAU48227

seq_documentation_block:

LOCUS TAU48227 2055 bp mRNA PLN 12-JUN-1996
 DEFINITION Triticum aestivum soluble starch synthase mRNA, partial cds.
 ACCESSION U48227
 VERSION U48227.1 GI:1373149

KEYWORDS

SOURCE wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 2055)
 AUTHORS Block,M., Loerz,H. and Luetticke,S.
 TITLE Direct Submission
 JOURNAL Submitted (01-FEB-1996) Martina Block, University of Hamburg,
 Institute of General Botany, Centre of Applied Molecular Biology,
 AMP II, Ohnhorstst. 18, Hamburg, 22609, Germany

COMMENT On Jun 12, 1996 this sequence version replaced gi:1335887.

FEATURES

source

1..2055

/organism="Triticum aestivum"

/cultivar="T.a., Florida"

/db_xref="taxon:4565"

/tissue_type="endosperm of kernels"

<1..1834

/codon_start=2

/product="soluble starch synthase"

/protein_id="AAB02197.1"

/db_xref="GI:1373150"

/translation="EAAPYAKSGGLGVCSLPIALAAARGHRVMVMPRLNGSSDKN

YAKALYAKHIKIPCFGSGSEVTFEYRDNDVDFVDPHSVHRPGSLYCDNFAGFG

NQRYTLTYACEAPLILELGGYIYQNCMFVNDWHSALVPLLAAXYRPYGVYRD

SRSTLVHNLHAGQVEPASTYDGLPPEWYFPEWARRHALDKGEAVNFKG

AVVTADRTVSQGYSWEVTTAEGGGLNELLSRSKVLNGIVNGIDINDWPTDKC

LPHYSVDDLSGKAKAELOKELGLPVREDVPLIGFGLDYQKSIDLILKAIPELM

REDVQFVMLSGSDPIFECHMRSTESSYKDKFGWGVFVPSHRITAGCDILMPSRF

EPGGLNQLYAMQYGVVPHVHGTGLRDITETNPFKAGEEGTGWAFSLPTVDMKLA

LRYAMSTFREHKFSEWGLMKRMGTWDHAPSSSRSSSSSSGNTNPTSCRRGLGRS

KCESPALKITSSSSFGPEGYPTCLRPATVESCACLLFWAGSRTYDGC AAAAATAS

GGRLQFQIRKGKCAAGWLTAHHSDGSLSVRVTAIRNQLVTL"

BASE COUNT 570 a 399 c 534 g 552 t

ORIGIN

alignment_scores:

Quality: 3313.00 Length: 610

Ratio: 5.431 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-674-824-2 x TAU48227

Align seg 1/1 to: TAU48227 from: 1 to: 2055

147 GluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySe 163

2 GAAGCTGCTCTTATGCAAGTCAAGGGGGTGGAGATGTTGTGTTTC 51

163 rLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMetp 180

52 GTTACCAATTGCTCTTCTGCTGCTGCTCACCAGTGTGTTGTTATGC 101

180 roArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyr 196

102 CAAGATACCTTAATGGGTCTCTGATAAAACTATGCAAGGATTTATAC 151

197 ThrAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValTh 213

152 ACTGCGAAGCACATTAAGATTCCATGCTTTGGGGGATCAGATGATGAC 201

213 rPhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHisP 230

|||||
202 CTTTTTTCATGAGTATAGACAACTCGATTGGGTGTTGTCGATCAATC 251
230 roSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPhe 246
252 CGTCATATACAGACACAGGAGTTTATATGGAGATAATTTGGTGCTTT 301
247 GlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAl 263
302 GGTGATAATCAGTTCAGATACACACTCCTTTGCTATGCTGCATCGAGGC 351
263 aProLeuIleuLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetP 280
352 CCCACTAATCCTGAAATGGGAGGATATATTATGGACAAATGCATGT 401
280 heValValAsnAspTrpHisAlaSerLeuValProValLeuLeuAlaAla 296
402 TTGTTGTGAAGATTGGCATGCCAGCCTTGTGCCAGTCTTCTTGCTGCA 451
297 LysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLeuValII 313
452 AAATATAGACCATACGGTGTGTACAGAGATTCCCGCAGCACCTTGTAT 501
313 eHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrProAspL 330
502 ACATAATTTAGACATACAGGTGTGGAGCTGCAAGTACATATCTCTGATC 551
330 euGlyLeuProGluTrpTyrGlyAlaLeuGluTrpValPheProGlu 346
552 TGGGATTCCTCCTGAATGGTATGAGGCTTTAGAATGGGTATTTCCAGAA 601
347 TrpAlaArgArgHisAlaLeuAspGlyGlyGluAlaValAsnPheLeuLy 363
602 TGGGCAAGGAGGCATGCCCTTGACAAGGTGAGGAGTTAACTTTTGA 651
363 sGlyAlaValValThrAlaAspArgIleValThrValSerGlnGlyTyrS 380
652 AGGACCACTGTGTACAGCAGATCGGATGTGACCGTCAGTCAGGGTTATT 701
380 erTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeu 396
702 CATGGAGGTTCACAACTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTA 751
397 SerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAs 413
752 AGCTCCCGAAAAAGTGATTGAATGGAATTTGAATGGAATTGACATTA 801
413 nAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrSerVala 430
802 TGATTTGGAACCCACACAGCAAGTGTCTCCCTCATCATATTCTGTGG 851
430 spAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeu 446
852 ATGACCTCTCTGGAAGGCCAANTGTAAAGCTGAATTCAGAAGGAGTTG 901
447 GlyLeuProValArgGluAspValProLeuIleGlyPheIleGlyArgLe 463
902 GGTTTACCTGTAAGGAGGAGTGTCTCTGATTTGCTTTATTGGAAGACT 951
463 uAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProGluLeuM 480
952 GGATTTACCAAGAGAGGATTTGATCTCATTTAANTGCGCATTCAGAGCTCA 1001
480 eArgGluAspValGlnPheValMetLeuGlySerGlyAspProIlePhe 496
1002 TGAGGAGGAGCGTGCAATTTGTCATGCTTGGATCTGGGATCCAATTTT 1051
497 GluGlyTrpMetArgSerThrGluSerSerTyrLysAspLysPheArgG 513
1052 GAAGGCTGGATGAGATCTACCGAGTCGATTAAGATAAATTCCTGG 1101
513 yTrpValGlyPheSerValProValSerHisArgIleThrAlaGlyCysA 530
|||||

1102 ATGGGTTGGATTTAGTGTTCAGTTTCCACAGAAATAACTGCAGGTTGGC 1151
530 spIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeu 546
1152 ATATATTGTTAATGCCATCGAGATTTGAACCTTGGGTCTTTAATCAGCTA 1201
547 TyrAlaMetGlnTyrGlyThrValProValValHisGlyThrGlyGlyLe 563
1202 TATGCTATGCAATATGGTACAGTTCTCTGTAGTTTCATGAACCTGGGGCCT 1251
563 uArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyGluGluG 580
1252 CCGAGACACAGTCGAGACCTTCAACCTTTTGGTGCAAAAGAGAGGAGG 1301
580 LyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLeuTrpAla 596
1302 GTACAGGTGGGCTTCTCACCGCTTAACCGTGAGCAAGATGTTGTGGCA 1351
597 LeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGluGl 613
1352 TTGCGAACCGCATGTCGACATTCAGGGAGCACAAAGCCGTCCTGGGAGG 1401
613 yLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAlaProS 630
1402 GCTCATGAAGCGAGGCATGACGAAAGACCATACGTGGGACCATGCCCGCA 1451
630 exSerThrSerArgSerSerGlyProSerTrpThrAsnProThrSer 646
1452 GCAGTACGAGCAGATCTTCGAGTGGGCTTCGTGGACCAACCCCTACGTCA 1501
647 CysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaLeuLy 663
1502 TGTAGACGGGACTGGGGAGGTCCAAGTCGAGTCGAGTCTCCTTCAGCTGAA 1551
663 sthrSerSerSerPheArgGlyProGluGlyTyrProCysThrLeuA 680
1552 GACATCTCTTCATCTTCCGGGCCCGGAAGGATACCCCTGTACATTGC 1601
680 rgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAla 696
1602 GTTGTCTCTACAGTAGAGTCGAATCGCCCTGCTTGTGTTGGTTCGCC 1651
697 GlySerArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerGl 713
1652 GGTTCGAGAACATATGACGCTCTGCTGCTGCGCGGTGACAGCTTCGGG 1701
713 yGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyT 730
1702 TGGACGACAGTTACAGTTTGGGGAATAAGAAAGGATGTGCTGCAGGAT 1751
730 rPLeuThrAlaLysHisSerAspGlySerLeuSerValArgValThr 746
1752 GGTAAACAGCAAGCACCACTCAGATGGCAGCCTCTCTGTCGCTGTACA 1801
747 AlaGluIleArgAsnGlnLeuValThrLeu 756
1802 GCTCAATCAGAAACCAACTGGTGACTCTT 1831
seq_name: gb_pl:AF165890
seq_documentation_block: 2585 bp mRNA PLN 17-AUG-1999
LOCUS AF165890
DEFINITION Oryza sativa subsp. japonica soluble starch synthase mRNA, complete cds.
ACCESSION AF165890
VERSION AF165890.1 GI:5734102
KEYWORDS
SOURCE Oryza sativa subsp. japonica.
ORGANISM Oryza sativa subsp. japonica.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 2585)

AUTHORS Junwang, X. and Zhen, Z.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1999) Group 601, Genetics Institute of CAS, Datun Road, Beijing, Beijing 100101, China
FEATURES Location/Qualifiers

```

location/qualifiers
1..2585
/organism="Oryza sativa subsp. japonica"
/cultivar="Nanjing37"
/sub_species="japonica"
/db_xref="taxon:39947"
139..2019
CDS

```

```

/codon_start=1
/product="soluble starch synthase"
/protein_id="AAD09850.1"
/db_xref="GI:5734103"
/translation="MATAAGMIGAACLVAPQVRGRRLRQVRRCVAAELSRDGS
AQRPDLAPLVQPLVPTFLVPTTTPAQTOSPAPATPPPLPSGVGEIPEDEGLT
EDSIDKTIIVASEQSEIINDVKEAQAKVTSVVFVTEASPYAKSGGLDVGCSLPI
ALALGRHVWVWPMYMGALNKNFANAFYTEKHKIKPCFGEHEVTFHFYIEDSDW
VFDHSPYRGNLQYDNGFAGDQNFRTLLCYAAACEAPLILEGYYIYQCKMFV
NDHNASLVPVLAARYGYVDVARSVLVTHNLAHQCEVAPETPDLGTPCMWYCAL
EWFPFWARHIALDKGEAVNFKLGAVATADRVTVSQCYSEWNTTASGGGGLNELSS
KSVLAVNGINDIDINDWNPSTDKPLPYHVVDLISGRACKAELQKELGPIRPDVP
IGFTIGRLYQKIGIDILIKLAIPLMDRNTQIFVNLGSDGPGEGMWRSTESYGRDKFRG
VGFSPVSRVHTAGDILLKPNRVEPFCGNQLYAQVTPVHVHGTGKLDRTYENENP
FAEGQGGTGWAFSPVLTLEKNRAGTNDNFQIQTGVLLAGSNEARHVKRLYMGPCR

```

BASE COUNT	660 a	568 c	706 g	651 t
ORIGIN				

alignment_scores:		
Quality:	2829.50	Length: 726
Ratio:	4.527	Gaps: 18
Percent Similarity:	86.088	Percent Identity: 76.446

alignment_block:
US-09-674-824-2 x AF165890

Align seg 1/1 to: AF165890 from: 1 to: 2585

```

2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
   |||::  |||::|||::|||::|||::  |||::|||::|||::|||::
151 GCGGGGATGGGGATGGGGGGCGCTGCCTGCTGCGCGCCGAGGTGAGG.. 198

```

```

17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
    |||:::
199 .....CCGGGGAGG.....AGGT 211

```

34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||:::||||::: ||| ||| |||||||||||||
212 TCGGGCTCCAGCGGGTGCAGAG...CGGTGCGTGGCGGAGCTGAGCAGG 258

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaPro..... 65
:::| | :|||
259 GACGGTGGGTCTCGGC.....CAGCGCCCCGCTGGCACCGCGGCC 298

66ProLeuValProGlyPheLeuAlaPro.....Prop 76
 |||:||||| |||:|||||
 297 GCTGGTGAAGCAGCCGGTCTCGCGACCTCTCTGTCGCGACGTCGACGC 346

76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
|||||:::||||| |||||
347 CACC CGGCCACAGCATGTCCCGGGGCCCGGCACCGCGCGCTTG 396

91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyII 107
 |||||:|||||:|||||
 397 CGGGACTCCGGCGTGGGGAGATCGAGCCCGAT.....CTAGAAGGTCT 440

107 eAluGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAsps 124
 :::::::::::::::::::: ::::::::::::::::::::
 441 CACAGAAGATCCATCGACAAACAATTTTGTGGCTACTGACGAGACT 490

124 erGluleMetAspAlaasnGluInProGlnAlaLysValThrArgSer 140
125 :
491 CTGAGATCATGGTGTGAAGGCAAGCTCAAAGCTAAAGTAACGCCAGC 540
141 IleValPheValThrclyGluAlaAlaProTyrrAlaLysSerGlycyle 157
142 :
541 GTTGCTTTGTAAACCGGTGAAGCTTCTCTTATGCAAAAGTCAGGTGGACT 590
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaargGlyHisA 174
158 :
591 AGGAGATGTTTGGTTCACTGCCAATTGCCTCTCTCTCGTSGTCATC 640
174 rgValMetValValMetProArqTyrrLeuAsnGlySerSerAspLysAsn 190
175 :
641 GTGTGATGGTTGTAATGCCGAGATACATGAACGGGGCGCTTGAACAAAAT 690
191 TyrAlaLysAlaLeuTyrrThrAlaLysHisIleLysIleProCysPheGI 207
192 :
691 TTTTGCAAACGCCATTTTACACTGAGAAGCACATTAAAGATTCCATGCTTTGG 740
207 yGlySerHisGluValThrPhePheHisGluTyrrArgAspAsnValAspt 224
208 :
741 CGGAGAACATGAAGTTACTTTTTTTCACGAGTAGGGAATCTCTGTGATT 790
224 rpValPheValAspHisProSerTyrrHisArqProGlySerLeuTyrrcly 240
225 :
791 GGGTCTTTGTGATCATCCCTCATATCATAGACCTGGNAATTTGTATGGA 840
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrrThrLeuLeuCy 257
242 :
841 GATAATTTTGGTGTCTTTTGGCGGATAATCAGTTCAGATACACACTCCTGTG 890
257 stYrAlaAlaCysGluAlaProLeuIleLeuLeuLeuClyGlyTyrrilet 274
275 :
891 CTATCGGCGGTGTGAGCCCCCATTAANTCTTTGAACCTGGGAGGATATATCT 940
274 yrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
275 :
941 ATGGACAGAAATGCATGTCTTGTGTGAATGATTGGCATGCCAGTCTGTGTG 990
291 ProValLeuLeuAlaAlaLysTyrrArgProTyrrclyValtyrrArgaspSe 307
292 :
991 CCAGTCTCTTGTGTCGCAAAATATAGACCATATGGTGTTTACAGGGATGC 1040
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
308 :
1041 CCGCAGTGTCTTGTTCATACATAATCTAGCACATCAGGGTGTGGAGCCTG 1090
324 laSerThrTyrrProAspleuGlyLeuProProGluTrpTyrrGlyAlaLeu 340
325 :
1091 CCAGTACATAPCCATPGACCTGGGATTTGCCACTGAATGTGATGAGCATTA 1140
341 GluTrpValPheProGluTrpAlaArqArgHisAlaLeuAspLysGlyGL 357
1141 GAATGGGTGTTCACAGAGTGGCAAGCGGCATGCCCTTGACAAGGGTGA 1190
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgileValTr 374
1191 GGCACTCAATTTTTTAAAGCGCGCAGTGTGTGACAGCAGATTCGAATTTGTGA 1240
374 hrValSerGinGlyTyrrSerTrpGluValThrThrAlaGluGlyGlyGln 390
1241 CTGTACGCCAGGGGTATTCTATGGGAGGTCACAACCTGCTGAAGGTGGGC AA 1290
391 GlyLeuAsnGluLeuLeuSerSerArqLysSerValLeuAsnGlyIleVa 407
1291 GGCTCTCAATCAGCTCTTAAGCTCCCGGAAGAGTGTTATTGAATGGAATTC 1340
407 lAsnGlyIleAspileAsnAspTrpAsnProThrThrAspLysCysLeuP 424
1341 AATGGNATTCACATTAATGATTGGAAACCCATCCACAGACAAGTTTCTCC 1390
424 roHishistyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440

```

seq_documentation_block: 2533 bp      RNA      PAT      29-SEP-1997
LOCUS      E06904
DEFINITION      cDNA encoding soluble starch synthase.
ACCESSION      E06904
VERSION      E06904.1 GI:2175060
KEYWORDS      JP 1994070779-A/1.
SOURCE      Oryza sativa.
ORGANISM      Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
Baba, T. and Shimada, H.
SOLUBLE RICE STARCH SYNTHETASE GENE AND ITS USE
Patent: JP 1994070779-A 1 15-MAR-1994;
MITSUI GIYOUSAI SHOKUBUTSU BIO KENKYUSHO:KK
OS      Oryza sativa
PN      JP 1994070779-A/1
PD      15-MAR-1994
PF      07-JUL-1992 JP 1992179947
PI      BABA TADASHI, SHIMADA HIROAKI
PC      C12N15/52,A01H5/00;
CC      strandedness: Single;
CC      topology: Linear;
CC      hypothetical: No;
CC      anti-sense: No;
CC      *source: tissue_type=ripening seed;
FH      key      Location/Qualifiers
FT      5'UTR      1..113
FT      CDS      114..1994
FT      /product='Soluble starch synthase' FT
transit_peptide 114..452
FT      3'UTR      1992..2533.
FEATURES
source      Location/Qualifiers
1..2533
/organism="Oryza sativa"
/db_xref="taxon:4530"
BASE COUNT      630 a 550 c 701 g 652 t
ORIGIN

alignment_scores:
Quality: 2821.50      Length: 726
Ratio: 4.522      Gaps: 18
Percent Similarity: 85.950      Percent Identity: 76.309

alignment_block:
US-09-674-824-2 x E06904

Align seg 1/1 to: E06904 from: 1 to: 2533

2 AlaaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
|||||  |||:|||||:|||||:|||||  |||:|||||
126 GCGGGATGGGATCGGGCGCGTGCCTGGTGGCGCCGAGTGAGG... 173
|||||  |||:|||||:|||||:|||||  |||:|||||
17 uArgAlaAspProAlaThrAlaAlaAlaArgAlaSerAlaCysValValArgA 34
|||||  |||:|||||:|||||:|||||  |||:|||||
174 .....CCGGGGAGG.....AGGT 186
|||||  |||:|||||:|||||:|||||  |||:|||||
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||  |||:|||||:|||||:|||||  |||:|||||
187 TGGCGGTCCAGCGGGTGGCGAGG...CGGTGGCTGGCGAGCTGAGCAGG 233
|||||  |||:|||||:|||||:|||||  |||:|||||
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnAlaPro..... 65
|||||  |||:|||||:|||||:|||||  |||:|||||
234 GACGGTGGGTGGCGG.....CAGGCGCGCTGGCACCAGCGGCC 271
|||||  |||:|||||:|||||:|||||  |||:|||||
66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76
|||||  |||:|||||:|||||:|||||  |||:|||||
272 GCTGTGTGAAGCAGCGGTCTGCCACCTTCCTCGTCCGAGCTGACGCG 321
|||||  |||:|||||:|||||:|||||  |||:|||||
76 rbcraAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
|||||  |||:|||||:|||||:|||||  |||:|||||

```


672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
 ::::::::::::::::::::|
 2166 TCAAGGTACCCCTTGTAGG.....CCTGCTACTTGCCCA..... 2199

689 ysAlaCysLeuLeuTrrPhe 695

2200ATATGGTTC 2208

seq_name: gb_pl:RICSSS

seq_documentation_block: 2533 bp mRNA PLN 01-FEB-2000
 LOCUS RICSSS
 DEFINITION Rice mRNA for soluble starch synthase, complete cds.
 ACCESSION D16202
 VERSION D16202.1 GI:450484
 KEYWORDS soluble starch synthase.
 SOURCE Oryza sativa Immature seed, cDNA to mRNA, clones RS[1, 2, 3 and 4].
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoae; Oryza.

REFERENCE 1 (bases 1 to 2533)
 AUTHORS Baba,T., Nishihara,M., Mizuno,K., Kawasaki,T., Shimada,H.,
 Kobayashi,E., Ohnishi,S., Tanaka,K. and Arai,Y.
 IDENTIFICATION, cDNA cloning, and gene expression of soluble starch
 synthase in rice (Oryza sativa L.) immature seeds
 TITLE Plant Physiol. 103 (2), 565-573 (1993)
 JOURNAL 94302151
 MEDLINE 2 (bases 1 to 2533)
 REFERENCE 2 (bases 1 to 2533)
 AUTHORS Baba,T.
 TITLE Direct Submission

JOURNAL Submitted (06-MAY-1993) to the DDBJ/EMBL/GenBank databases, Tadashi
 Baba, University of Tsukuba, Institute of Applied Biochemistry,
 Tennohdai 1-1-1, Tsukuba Science City, Ibaraki 305, Japan
 (Tel:298-53-6632, Fax:298-53-6632)
 COMMENT Submitted (06-May-1993) to DDBJ by:
 Tadashi Baba
 Institute of Applied Biochemistry
 University of Tsukuba
 Tennohdai 1-1-1
 Tsukuba Science City
 Ibaraki 305
 Japan
 Phone: 0298-53-6632
 Fax: 0298-53-6632.

FEATURES

source

Location/Qualifiers

1..2533

/organism="Oryza sativa"

/db_xref="taxon:4530"

/dev_stage="Immature seed"

1..113

/codon_start=1

/product="soluble starch synthase precursor"

/protein_id="BAA03739.1"

/db_xref="GI:450485"

114..1994

transit_peptide 114..452

CDS

5'UTR

mat_peptide 453..1991

3'UTR

product="soluble starch synthase"

1995..2533

polya_signal 2487..2492
 BASE COUNT 630 a 550 c 701 g 652 t
 ORIGIN
 alignment_scores:
 Quality: 2821.50 Length: 726
 Ratio: 4.522 Gaps: 18
 Percent Similarity: 85.950 Percent Identity: 76.309

alignment_block:

US-09-674-824-2 x RICSSS ..

Align seg 1/1 to: RICSSS from: 1 to: 2533

2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 126 GCGGGATGGGATCGGGCGGGCGGCTGCTGGTGGCGCGGAGTGAGG... 173
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 174CCGGGAGG.....AGGT 186
 34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 187 TGC GGCTCCAGCGGGTCCGGAGG...CGGTGGTGGCGGAGCTGAGCAGG 233
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 234 GACGGTGGTGGCG.....CACGGCGCGTGGCACCGGGGCC 271
 66ProLeuValProGlyPheLeuAlaPro.....Prop 76
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 272 GCTGGTGAAGCAGCGGCTCTCGCGACCTTCCTCGTCCGACGTCGACGC 321
 76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 322 CACCGCGCGCCACGAGTCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 371
 91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyI 107
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 372 CCGGACTCCGCGGTGGGGAGATCGAGCCGAT.....CTAGAAGGTCT 415
 107 eAlaGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAsp 124
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 416 CACAGAAGATCCATCGACAAACAATTTTGTGGCTAGTAGGAGGAGT 465
 124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSer 140
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 466 CTGAGATCATGTGTGAGGAGGCAAGCTCAAGCTAAGTAACACGACG 515
 141 IleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 516 GTTGCTTTGTAACCGGTGAAGCTTCTCTTATGCAAGTCAGGTGGACT 565
 157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHis 174
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 566 AGGAGATGTTGTGGTTTCACTGCCAATTTGCTTCTCTCTCGTGTCTATC 615
 174 rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 616 GTGTGATGTTGTAATCGGAGATACATGAACGGGCGCTTGAACAAAAT 665
 191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 666 TTGTCAAACGCAATTTTACACTGAGAAGCACATTAAGATTCCTGCTTGG 715
 207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAsp 224
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 716 CGGAGAACATGAATTTACTTTTTTCACGAGTATAGGATTCGTGTGATT 765
 224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||


```
766 GGGTGTGTTTGTGATCCCTCATATCATAGACCTGGAAATTTGTATGA 815
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
816 GATAATTTTGGTGTGTTTGGCGATAATCAGTTCAGATACACACTCCTGTG 865
257 sTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleT 274
866 CTATGGCGGTGTGAAGCCCATTAATTTCTTGAAGTGGAGGATATATCT 915
274 yGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
916 ATGSGACAAAATGCGATGTTTGTGTAATGATTGGCATGCCAGTCTTGTG 965
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
966 CCAGTCTCTTGTGTCGCAAAATATAGACCATATGTTGTTCACAGGGATGC 1015
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
1016 CCGCAGTCTTCTGTGCATACATAATCTAGCACATCAGGGTGTGGAGCCTG 1065
324 laSerThrTyrProAspLeuGlyLeuProGluTyrTrpTyrGlyAlaLeu 340
1066 CCAGTACATATCCAGACTGGGATGTCACCTGAATGATGGAGCATTA 1115
341 GluTrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyG1 357
1116 GAATGGGTGTTTCCAGAGTGGCAGGCGCATGCCCTTGACAAGGGTGA 1165
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
1166 GGCAGTCAATTTTAAAGGCGCAGTTGTGACAGCATCGAATTTGTGA 1215
374 hrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGln 390
1216 CTGTGCCAGCGGTATTCATGGGAGGTGCACACTGCTGAAGTGGGCAA 1265
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
1266 GGCTCAATGAGCTCTTAAAGCTCCGGAAGAGTGTATTGAATGAATTTGT 1315
407 laSnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424
1316 AAATGGAAATGACATTAATGATTGGAACCATCCACAGACAAGTTCTCC 1365
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
1366 CTTATCATTTATCTGTTGATGACCTGTCGGAAGGCAAGCTGTAAAGCT 1415
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI1 457
1416 GAATTGCGAAGGAGCTGGGTGTTTACCTATTAAGGCCGATGTCCTCTGAT 1465
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
1466 TGGCTTTATTTGGAAGATGACATATCAAAAGGCATGATGATTAATTAAC 1515
474 etAlaIleProGluLeuMetArgLysValGlnPheValMetLeuGly 490
1516 TTGCCATTCAGATCTCATCGGGAACAATATTCAATTCGTGTCATGCTTGA 1565
491 SerGlyAspProIlePheGluGlyTyrPheMetArgSerThrGluSerSerTy 507
1566 TCTGGTGACCCAGGTTTGAAGGATGGATGAGATCCACAGAATCAGGGTA 1615
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524
1616 CAGGATAAATTTTCGTGGATGGGTGGAATTAAGTTTCCAGTTTCCACC 1665
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540
1666 GAATAACTGAGGTTGCGATATATTTGATGTCATCCATCCAGATTCCGAACCT 1715
```

```
541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557
1716 TGTGGCCTCAATCAGCTATATGCTATATCAATATGTTACAGTGGCTGTGT 1765
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574
1766 TCATGGAACTGGAGGCTCAGAGATACAGTGGAGAATTTTAAACCGTTTG 1815
574 lyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590
1816 CTGGAAGAAGAGAGAGAGGTCAGAGGTGGGCAATTCGCCCACTAACCAT 1865
591 AspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607
1866 GAAAAAATGCTGTGGGCAATTCGGGATGGCAATTTGCACATACAGGGAAC 1915
607 isLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
1916 ACAAGTCTCTTGGGAGGGTCTAATGAAGCGAGGCATGTCAAGCGACTTT 1965
624 ThrTrpAspHisAlaProSerSerThrSerArgSerSerGlyProS 640
1966 ACATGGGACCATCCCGCTCACAGTATGACAGATCTTCAATGGGCCCTT 2015
640 erTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLys.. 655
2016 CATGGATCAACCATATGTCATGTAATGATTTGAAGGAGCAGCGCAATT 2065
656 CysGluSerPro.....SerAlaLeuLysThrS 665
2066 TCTCCAGGACCTCAATCTCTGCTTTCATGAGCGGAATGAAAACTT 2115
665 erSerSerPheArgGly.....Pr 672
2116 TGTACACTACATGGAAGGGAACCACTTATGCAAGTTCCAAACGATCAC 2165
672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
2166 TCAAGGTTACCTTTGAGG.....CCTGCTACTTGGCCA..... 2199
689 ysAlaCysLeuLeuTrpPhe 695
2200 .....ATATGGTTC 2208
seq_name: gb_pat:A93359
seq_documentation_block:
LOCUS A93359 2383 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9744472.
ACCESSION A93359
VERSION A93359.1 GI:6741623
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 2383)
Kossmann, J. and Froberg, C.
NUCLEIC ACID MOLECULES CODING SOLUBLE MAIZE STARCH SYNTHASES
Patent: WO 9744472-A 1 27-NOV-1997;
KOSSMANN JENS (DE); FROBERG CLAUS (DE)
FEATURES
source
1..2383
/organism="Zea mays"
/db_xref="taxon:4577"
/tissue_type="ENDOSPERM"
<2..1951
/function="STARCH SYNTHESIS"
/codon_start=1
/product="SOLUBLE STARCH SYNTHASE"
/protein_id="CAB69545.1"
```

```
/db_xref="GI:6741624"
/translation="ARGLLSISAMATPSAVGAACLLLARAAMPAAVGDRAPRRLQVR
LRRCVAELSRGPAWRPMPALLAPPLVPGFLAPAEPTGEPALTTPVPDAGLGLV
GVPEGIAEGSIDNTVVVASEQDSIIVGKEQAKVQNIVFVTGEASPVAKSGGLG
DYVGLSPVALAARGVHMVMPRYLNGTSDKNYANAFYTERHIRPCFGRHEVTFPH
EYRDSVDFVDPHSYHRPGLNGYDKFGAGDNQRYTLCCYAAACEAPLVLELGGYIY
GQNCMFVNDHMSLVPLVLLAAKYPYGVYKDSRSLIVHNLAHQGVPEASTYPDLGL
PEWYGALEMWPEWARHRLDKGEAVNFKGAVTADRIVTVSKGYSWEVTTABGGQ
GLNELSSRKSLVNLGVNGLVINGIDINDNPATDKICPCHYSVDDLSGKAKCKGALOKELG
PIRPDVLPIGLGRDYLQKIGDILQILIPDLMDREDYQVFMVLSGSGPELEDNMRSTESI
FKDKFRGVGFSPVSHRITAGCDLILMPSREFPCGLNQLYAMQIGTVPVHVATGCLR
DVENFEPGEGEQTGWAFPLATENNLMWTLRTAISTYREHKSSWELMKRGMKSD
FTWDHAAEQEYIOWAFIDRPVVM"
BASE COUNT      582 a      546 c      633 g      622 t
ORIGIN
```

```
alignment_scores:
  Quality: 2809.50      Length: 749
  Ratio: 4.411      Gaps: 10
  Percent Similarity: 85.047      Percent Identity: 73.565

alignment_block:
US-09-674-824-2 x A93359  ..

Align seg 1/1 to: A93359 from: 1 to: 2383

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
  ::::::::::::::::::::
41 TCGGCGGTGGGCGCGCGCTGCTCCTC..... 67

20 pProAlaThrAlaAlaAlaArgAlaSerAlaCysValValArgAla..... 34
  ::::::::::::::::::::

68 .CTCGCGGCGCGCGCTGCGCGCGCGCGCTGCGCGCGCGCGCGCGCG 116

35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
  ::::::::::::::::::::

117 GCGCGCTCCAGCGCGTGTGCGCGCGCGCGTGTGCGCGCGCGTGTGCGCG 166

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnGlnGlnGlnGln 67
  ::::::::::::::::::::

167 GAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 216

67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
  ::::::::::::::::::::

217 CTGCGCGCGCTTCTCTCGG...CCGCGCGCGCGCGCGCGCGCGCGCGG 263

84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
  ::::::::::::::::::::

264 CATTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 313

101 AspLeuLeuLeuGluGlyTleAlaGluAspSerIleAspSerIleIleVa 117
  ::::::::::::::::::::

314 GAA.....CTGAGAGGATGCTGAGAGTTCATCGATAACACAGTAGT 357

117 laAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
  ::::::::::::::::::::

358 TGTGGCAAGTGACCAAGATCTGAGATTGTGGTTGGAAGGAGCAAGCTC 407

134 laAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
  ::::::::::::::::::::

408 GAGCTAAAGTACACAAAACATTGCTTGTAACTGGCGAAGCTTCTCCT 457

151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
  ::::::::::::::::::::

458 TATGCAAGCTCTGGGGTCTAGAGATGTTGTGGTTTCATTGCCAGTTGC 507

167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
  ::::::::::::::::::::

508 TCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
  ::::::::::::::::::::

184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
  ::::::::::::::::::::
```

```
558 ATGGTACCTCCGATAAGAAATTATGCAAAATGCATTTTACACAGAAAACAC 607

201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
  ::::::::::::::::::::

608 ATTTCGGATTCATGCTTGGCGGTGAACATGAAGATTACCTTCTTCATGCA 657

217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
  ::::::::::::::::::::

658 GTATAGAGATTCAGTTGACTGGGTGTTGTTGATCATCCTCATATCACA 707

234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
  ::::::::::::::::::::

708 GACCTGGAAAATTTATATGGAGATTAAGTTGGTCTCTTTGGTGATAATCAG 757

251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
  ::::::::::::::::::::

758 TTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGGTCT 807

267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
  ::::::::::::::::::::

808 TGAATTGGGAGGATATATTATGACAGAAATTCATGTTGTTGTTCAATG 857

284 spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300
  ::::::::::::::::::::

858 ATTGGCATGCCAGCTCTAGTCCAGTCTCTTCTGCTGCANAAATATAGACCA 907

301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
  ::::::::::::::::::::

908 TATGGTCTTTTATAAAGACTCCCGCAGCATCTCTGTAATACATAATTTAGC 957

317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
  ::::::::::::::::::::

958 ACATCAGGGGTGAGAGCCTGCAAGCACATATCCTTGACCTTGGGTGGCCAC 1007

334 roGluTyrGlyValLeuGluTyrValPheProGluTyrAlaArgArg 350
  ::::::::::::::::::::

1008 CTGAATGGTATGGAGCTCTGGAGTGGGTATTCCTGTAATGGCGAGGAGG 1057

351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
  ::::::::::::::::::::

1058 CATGCCCTTGACAAAGGTGAGGCGAGTTAAATTTTGAAGGTGCGAGTTGT 1107

367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
  ::::::::::::::::::::

1108 GACAGCAGATCGAATCGTACTGTCTAGTGAAGGTTATTCATGGGAGGTCA 1157

384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
  ::::::::::::::::::::

1158 CACTGCTGAAAGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAAG 1207

401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
  ::::::::::::::::::::

1208 AGTGATTTAAACGGAATTTGTAATGGAATTCACATTAATGATTGGAAACC 1257

417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
  ::::::::::::::::::::

1258 TGCCACACACAAATGTATCCCTCTCATATTCTGTTGATGACCTCTCTG 1307

434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
  ::::::::::::::::::::

1308 GAAAGGCCAAATGTAAGGTGCAATTCAGAGAGGAGCTGGGGTTTACCTATA 1357

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
  ::::::::::::::::::::

1358 AGGCTGATGTTCTCTGATGGCTTTTATGGAAGATTGGATTATCAGAA 1407

467 sgIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
  ::::::::::::::::::::

1408 AGGCATTGATCTCATCAACTATCATACCAGATCTCATGCGGGAAGATG 1457

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
  ::::::::::::::::::::

1458 TTCAATTTTCTCATGCTTGGTGTGACCCAGAGCTTGAAGATTGGATG 1507
```

501 ArgSerThrGluSerSerThrLysAspLysPheArgGlyTrpValGlyPh 517
|||||
1508 AGATCTACAGATCGATCTCAAGGATAAATTCGTGGATGGTTGGATT 1557
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1558 TAGTGTTCAGATTTCACCGAATAACTGCCGGTCGCGATATATTGTTAA 1607
534 eProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrrAlaMetGln 550
1608 TGGCATCCAGATTCGAACCTTGTGGTCAATCAGCTATATCTATGCGAG 1657
551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567
1658 TATGGCACAGTTCCTGTTGTCATGCAACTGGGGCCCTTAGACATACCGT 1707
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584
1708 GGAGAACTTCAACCTTTCGGTGAGAAATGGAGAGAGGAGGACAGGTGG 1757
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1758 CATTGCCACCTTACCACAGAAACATGTTGTGGACATTCGGAATGCA 1807
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1808 ATATCTACATACAGGAACACAAAGTCTCTCGGAAGGCGCTAATGAAGCG 1857
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
1858 AGGCATGTCAAAGAACTTCACGTTGGGACCATGCCCTGAACAATACGAG 1907
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgGly 650
1908 AAATCTTCAGTGGCCCTTCATCGATCGACCTATGTCATGTAAAGAG 1957
650 yLeuGlyArgSerLysCys.GluSerProSerAlaLeuLysThrSerSer 666
1958 A.....CCAAAGTGGTGGTCTCTTGAAGATCATCAGTTCATCATCC 1998
667 SerSerPheArgGlyProGlyGlyTrpPro.....CysThrLeuArgCy 681
1999 TATAGTAACTCAATGATGAAGAAGAACCCCTGTACATTCATGGAAGC 2048
681 sProLathrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlys 698
2049 AGACCGGCTATTGGCTCCATGCTCAATGCTGCTTGGCTGCT.... 2094
698 eArgThrTyrrAspGlyCysAlaAlaAlaAlaValThrAlaSerGlyGly 714
2095TGCCCTGATGACCGGATGCGAGTGAAGTCA 2127
715 ArgGlnLeuGlnPheTrpGlyLeuArgLysGlyCysAla.....Al 728
2128 GCCGAACGACAGTTTGAAGATAGGAGGAGGAGCTGGAACGACGTCAGC 2177
728 aGlyTrpLeuThrAlaLysHisSerAspGlySerLeuSer 742
2178 AGGCAGCCTCGCGCTGATTCATATGGAACAAGCTGGAGTCAGT 2220
seq_name: gb_pl:AF036891

seq_documentation_block:
LOCUS AF036891 2991 bp mRNA 03-MAY-2001
DEFINITION Zea mays starch synthase I precursor (Ssl) mRNA, nuclear gene
encoding plastid protein, complete cds.
ACCESSION AF036891
VERSION AF036891.1 GI:2828011
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 2991)
Knight,M.E., Harn,C., Lilley,C.E.R., Guan,H., Singletary,G.W.,
Mu-Forster,C., Wasserman,B.P. and Keeling,P.L.
Molecular cloning of starch synthase I from maize (w64) endosperm
and expression in Escherichia coli
Plant J. 14 (5), 613-622 (1998)
98340555
9675904
2 (bases 1 to 2991)
Knight,M.E., Harn,C., Lilley,C.E.R., Guan,H.P., Singletary,G.W.,
Mu-Forster,C., Wasserman,B.P. and Keeling,P.L.
Direct Submission
Submitted (03-DEC-1997) ExSeed Genetics, 1568 Food Science
Building, ISU, Ames, IA 50011, USA
FEATURES
Location/Qualifiers
1..2991
/organism="Zea mays"
/strain="w64"
/db_xref="taxon:4577"
/chromosome="9"
/map="9; between np1266 and wx1"
/tissue_type="endosperm"
1..2991
/gene="Ssl"
602..2470
/gene="Ssl"
/EC_number="2.4.1.21"
/codon_start=1
/product="starch synthase I precursor"
/protein_id="AA89957.1"
/db_xref="GI:2828012"
/translation="WATPSAVGAACLLARAAPAAVGDRAAPRLQRLRRRCVAEL
SREGAPRLPLPALLAPLPVGLAPPAETGEPASTPPVPDAGLGLDGLPEGIAE
GSIDNTVVASQDSEIYVYKQARAKVTQSIIVFTGASPVAKSGGLDVCGLPVA
LAARGHVMVMPRYLNGTSDKNYANAFVTHIRPCFGEHETFFHEYRDSVDW
FVDHPSYHRYGDKFGAGDNQFRVTLCYAACAPLLELGGYLYGONCMFVN
DMHASLVPVLLAAKYRKYVYKDSRSLVHNLHAGQVEPASTYPLDGLPPMYGAL
WYFPEARRHALDKGEANFVKGAVTADIRVIVSKYSVEVTTAEGGGLNELLSR
KSLNGIVNGIDINDWNPATDKICPHYSDVDDLKGAKCKGALQKELGLTRPDVPLI
GFIRLDYOKGIDLIPLIDREDVQVFMGSGKPELEDMRSTESIFKDKFRGW
GFSVPVSHRITAGCDLILMPSRFPCLNOLYAMQYTPVVVHATGTRDVTENFNP
GENGQGTGWAFAPLTENMEVDIANCNIYIQGTQVLLGRANEARHVRRLHVGPCR"
transit_peptide 602..718
mat_peptide 719..2467
BASE COUNT 750 a 660 c 804 g 776 t 1 others
ORIGIN
alignment_scores:
Quality: 2776.50 Length: 747
Ratio: 4.442 Gaps: 11
Percent Similarity: 83.668 Percent Identity: 73.762
alignment_block:
US-09-674-824-2 x AF036891 ..
Align seg 1/1 to: AF036891 from: 1 to: 2991
4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
614 TCGCCGTCGGCGCCGCTCCTC..... 640
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
641 CTCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 689
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||:|||||: ||| ||| |||||||||||||||

690 GCGCGCTCCAGCGCTGCTGGCCGCCGCTGCTGCGGAGCTGAGCAGG 739
 51 GluGlyProAlaIaAArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67
 740 GAGGGCCCGCGCGCGCTGCCAGCCCGGCTGCTGGCGCCCGCT 789
 67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
 790 CGTGGCCCGCTTCCTCGG...CCGCGCGGAGCCACGCGGTGAGCGG 836
 84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
 837 CATCGACGCGCGCGCTGCGCGAGCGCGCTGGGGAGCTCGCT... 883
 101 AspLeuLeuGluGlyLeuAlaGluAspSerIleAspSerIleIleVa 117
 884 ...CTCGAACCTGAAGGATTGCTGAAGTTCCATCGATACACAGTAGT 930
 117 lAlaIaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
 931 TCTGCCAAGTGAGCAAGATTCTGACATTGTGGTGGAAAGGAGCAAGCTC 980
 134 laAlaIysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
 981 GAGCTAAAGTAACAAAGCATTTGCTTTGTAACCGCGCAAGCTTCTCCT 1030
 151 TyrAlaIysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
 1031 TATGCNAAGTCTGGGGTCTAGACATGTTGTGGTTCTATGCCAGTTGC 1080
 167 aLeuAlaIaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
 1081 TCTTGTGCTGCTGGTCAACGCTGTATGCTTCTAATGCCAGATATTAA 1130
 184 snGlySerSerAspIysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
 1131 ATGGTACCTCCGATAGAATTATGCAATGCAATTTTACACAGAAACAC 1180
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
 1181 ATTCGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTCCATGA 1230
 217 uTyrArgAspAsnValAspIrpValPheValAspHisProSerTyrHisA 234
 1231 GTATAGAGATTACGTTGACTGGGTGTTTGTTCATCATCCTCATATACA 1280
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
 1281 GACCTGGAAATTTATATGAGATAGTTTGGTGTCTTGGTGCATATCAG 1330
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
 1331 TTCAGATACACACTCCTTTGCTATGCTGCTGCTGAGGCTCCTTTGATCCT 1380
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValAlaAsnA 284
 1381 TGAATTGGGAGGATATTTATGACAGAAATTCATGTTGTGTGTCATG 1430
 284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
 1431 ATTTGGCATGCCATCTAGTGCAGCTCTTCTGCTGCAAAATATAGACCA 1480
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
 1481 TATGGTGTATTAAGACTCCCGACGATCTTGTGTAATACATAATTTAGC 1530
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334
 1531 ACATCAGGTTGATAGCCTGCAAGCACATATCTGACCTTGGTGGTCCAC 1580
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
 1581 CTGAATGGTATGAGCTCTGAGTGGGTATTCCCTGAATGGCGGAGGAG 1630

351 HisAlaLeuAspTysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
 1631 CATGCCCTTGACAAGGCTGAGCAGTTAAATTTTGAAGGTCAGTTGT 1680
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
 1681 GACAGCAGATCGAATCGTACTGTCTAGTAAGGTTATTCTGGGAGGTCA 1730
 384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
 1731 CAACTGCTGAGGTGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1780
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
 1781 AGTGATTAAACGGAATTGTAATGGAATTGACATTAAATGATTGGAACCC 1830
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
 1831 TGCCACAGACAAATGTATCCCTCTCATATTCTGTTGATGACCTCTCTG 1880
 434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
 1881 GAAAGCCAAATGTAAGGTGCATTGTCAGAGGAGCTGGGTTTACCTATA 1930
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
 1931 AGGCTGATGTTCTCTCTGATTGGCTTTATTGGAAGTTGGATTATCAGAA 1980
 467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
 1981 AGGATTGATCTCACTCACTTATCATACCAGATCTCATGCGGGAAGATG 2030
 484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
 2031 TTCATTTGTCATGCTGGATCTGTGTGACCCAGAGCTTGAAGATTGGATG 2080
 501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
 2081 AGATCTACAGATCGATCTTCAAGGATAAATTTCTGGATGGGTGGATT 2130
 517 eserValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
 2131 TAGTGTCCAGTTTCCCAACCAATAAATGCGCGCTGCGATATATTGTAA 2180
 534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
 2181 TGCCATCCAGATTCCGACCTTGTGCTCTCAATCAGCTATATGCTATGCG 2230
 551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
 2231 TATGTCAGATTCCTTGTGTCACCACTGGGGCCTTAGAGATACCGT 2280
 567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
 2281 GGAGAACTTCAACCTTTTCGGTGAGATGGAGACGAGGCTACAGGTTGG 2330
 584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAl 600
 2331 CATTCGACCCCTTAACACAGAAACATGTTTGTGGACATTCGGAACCTG 2380
 600 aMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysA 617
 2381 AATATCTACATACAGGGAACAACTCTCTCTGCGAAGGCTAATGAAGC 2430
 617 rgGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSe 633
 2431 GAGGCATGTCAAAGACTTCAGTGGGACCATGCGCGCTGAACAATACGAA 2480
 633 rArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArg 649
 2481 CAATCTTCCAGTGGGCTTTCATCGATCGACCC...GATGTTCAATGGA 2527

224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
|||||
729 GGGTGTGTTGATCATCCCTCATATCACAGACCTGGAAATTTATATGGA 778
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
|||||
779 GATAGTTTGTGCTCTTTGGTGATTAATCAGTTTCAGATACAGGCTCCCTTG 828
257 sTyxAlaAlaCysGluAlaProLeuLeuLeuLeuGlyGlyTyrlet 274
|||||
829 CTATGCTCATGTGAGGCTCTTTGGTCTTGAATTTGGAGGATATATTT 878
274 yrGlyGlnAsnCysMetPheValValAsnAspTyrPheHisAlaSerLeuVal 290
|||||
879 ATGGACAGAATTCGATGTTTGTGATGATGGCATGGCAGTCTAGTG 928
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
|||||
929 CCAGTCCTCTTCGTCGCAAAATATAGACCATATGGTGTTTATAAGACTC 978
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
|||||
979 CCGCAGCATTCCTGTAATACATAATTTAGCACATCAGGGTGTAGAGCCTG 1028
324 IaSerThrTyrProAspLeuGlyLeuProGluTyrTyrGlyAlaLeu 340
|||||
1029 CAAGCACATATCCTGACCTTGGTGGTCCACCTCAATGGTATGGAGCTCTG 1078
341 GluTyrPheProGluTyrPheAlaArgArgHisAlaLeuAspLysGlyG 357
|||||
1079 GAGTGGTATTCCTGATGGGCAAGGAGGCATGCCCTTGACAGGGTGA 1128
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
|||||
1129 GGCAGTAAATTTTGAAGGGTGCAGTTGTGACAGCAGATCGAATTTGGA 1178
374 hrValSerGlnGlyTyrSerTyrGluValThrAlaGluGlyGlyGln 390
|||||
1179 CTGTCAAGTAAGGTTATTTCATGGAGGTCACAACTGCTCAAGGTGGACAG 1228
391 GlyLeuAsnGlnLeuLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
|||||
1229 GGTCTCAATGAGCTCTTAAGCTCCCGAAAGAGTGTATTAAACGGAATGT 1278
407 lAsnGlyIleAspIleAsnAspTyrAsnProThrThrAspLysCysLeuP 424
|||||
1279 AAATGGAATGACATTAAATGATGGAACTTCGCGACGACAAATGATACC 1328
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
|||||
1329 CTTGTCAATTTCTGTGATGACCTCTCGGAAAGCCAAATGTAAAGT 1378
441 GluLeuGlnLysGluLeuGlyLeuProValatGluAspValProLeuI 457
|||||
1379 GCATTGCAAGAGGCTGGGTTTACCTATAAGGCCTGAAGTTCCTCTGAT 1428
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
|||||
1429 TGGTTTATTTGGAAGATGGAATATCAGAAAGCATTTGATCTCATTCAC 1478
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490
|||||
1479 TTATCATACCACATCTCATCGGAGCAGCATTCATTTGTCATGCTTGA 1528
491 SerGlyAspProIlePheGlyTyrPheMetArgSerThrGluSerSery 507
|||||
1529 TCTGGTGACCCAGAGCTCGAAGACTGGATGAGATCTACAGAGTCGACTT 1578
507 rLysAspLysPheArgGlyTyrPheValGlyPheSerValProValSerHis 524
|||||
1579 CAAGGATAAATTTCTGTGATGGTGGATTTAGTGTTCAGATTTCCACACC 1628
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540

1629 GAATAACTGCCGCTCGATATATTGTTAATGCCATCCAGATTTCGAACCT 1678
541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557
|||||
1679 TGTGGTCTCAATCAGCTATATGCTATGTCAGTATGGCACAGTTCCTCTTGT 1728
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574
|||||
1729 CCATGCAACTGGGGCTTAGAGATACTGTGGAGAACTTCAACCCTTTCG 1778
574 lAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590
|||||
1779 GTGAGAATGGAGAACAGGTACAGGTGGCATTCGCCACCCCTAACCCAG 1828
591 AsplysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGlu 607
|||||
1829 GAAACATGTTTGGGACATTCGGAACCTGCANTTTCGACATACAGGAGC 1878
607 lLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
|||||
1879 ACAATCTTCTGGGAAGGCTCATGAAGAGGGGCATGTCAAAAGACTTC 1928
624 ThrTrpAspHisAlaproSerSerThrSerArgSerSerSerGlyProS 640
|||||
1929 ACGTGGGACCATGCCGCTGAACAATACGACAAATCTTCAGTGGGCTT 1978
640 erTyrThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLysCys 656
|||||
1979 CATCGATCGACCTTATGTCATGTAAGGGGAC..... 2010
657 GluSerProSerAlaLeuLysThr...SerSerSerSerPheArgGlyP 672
|||||
2011 CACAAGTGGTGTGCTCTTAAAGATCTTCAGTCTTCATCTTCACAGTGAA 2060
672 roGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGl 688
|||||
2061 GGATGAGAAAACCCCTGTACATTACAT..... 2087
688 nCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAspGlyCysA 705
|||||
2088GGAGGGGACC.....GGCTGTT 2106
705 lAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnPheTrpGly 721
|||||
2107 GGGTGGCTCCAAATGTCCAGGTAGCTTACGGAGCCTGCTGTTGGCTG 2156
722 lIleArgLysGlyCysAlaAlaGly 729
|||||
2157 GTT.....TGTCCTCCAAAGGA 2171
seq_name: gb_pat:AR106494
seq_documentation_block:
LOCUS AR106494 1752 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 12 from patent US 6107060.
ACCESSION AR106494
VERSION AR106494.1 GI:12821024
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1752)
AUTHORS Keeling, P. and Guan, H.
TITLE Starch encapsulation
JOURNAL patent: US 6107060-A 12 22-AUG-2000;
FEATURES
Location/Qualifiers
1..1752
source /organism="unknown"
BASE COUNT 433 a 378 c 470 g 471 t
ORIGIN

alignment_scores:

Quality: 2713.50 Length: 585
 Ratio: 4.916 Gaps: 3
 Percent Similarity: 94.359 Percent Identity: 86.325

alignment_block:

US-09-674-824-2 x AR106494 ..

Align seg 1/1 to: AR106494 from: 1 to: 1752

```

45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnI 61
|||||
4 GTCCGGAGCTGACAGAGGAGGCCCGCCGCGCTGCCACCCGC 53
|||||
61 nGlnLeuAlaProLeuValProGlyPheLeuAlaProProProA 78
|||||
54 GCTGCTGGCCCGCGCTGCTGCGCGCTTCCTCGCG...CCGCGCGCG 100
|||||
78 laProAlaGlnSerProAlaProThrGlnProProLeuProAspAlaGly 94
|||||
101 AGCCACGGGTGAGCCGGCATCGACGCCGCCGCCGCGCGCGCGCGC 150
|||||
95 ValGlyGluLeuAlaProAspLeuLeuGluGlyLeuAlaGluaspSe 111
|||||
151 CTGGGGACCTCGGT.....CTCGAACCTGAAGGGATTGCTGAAGGTT 194
|||||
111 rleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMetA 128
|||||
195 CATCGATAACACGACTAGTCTGGCAAGTGAGCAAGATTCTGAGATTGTGG 244
|||||
128 spAlaasnGluGlnProGlnAlaLysValThrArgSerIleValPheVal 144
|||||
245 TTGGAAAGGAGCAAGCTCGAGCTAAAGTAACACAAAGCATGTCTTTGTA 294
|||||
145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyLeuGlyLeuGlyValCy 161
|||||
295 ACCGGGAACTCTCTTATGCAAGCTCGGGGGTCTAGGAGATGTTG 344
|||||
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValy 178
|||||
345 TGGTTTATGCCAGTTGCTCTTCTGCTGCTGCTGCTACCGGTGATGGTTG 394
|||||
178 alMetProArgTyrLeuAsnGlySerSerAspLysasnTyrAlaLysAla 194
|||||
395 TAATGCCAGATATTAAATGATGATACCTCGCATTAAGAAATTATGCAAAATGCA 444
|||||
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisG1 211
|||||
445 TTTTACACAGAAACACATTCGGATTCATGCTTTGGCGGTGAACATGA 494
|||||
211 uValThrPheHisGluTyrArgAspAsnValAspTrpValPheValA 228
|||||
495 AGTTACCTCTTCCATGAGTATAGATTCAGTTGACTGGGTGTTGTTG 544
|||||
228 sphisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGly 244
|||||
545 ATCATCCCTCATATCACACCTCGGAATTTATATGGAGATAGTTTGGT 594
|||||
245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCy 261
|||||
595 GCTTTTGGTGATAATCAGTTTCAGATACACACTCCTTTGCTATGCTGCATG 644
|||||
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnC 278
|||||
645 TGAGGCTCCTTTGATCCTTGAATTTGGAGGATATATTTATGGACAGAAAT 694
|||||
278 ysmetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294
|||||
695 GCATGTTTGTTCATGATTGGCATGGCCAGTCTAGTGCACGCTCCTTCT 744
|||||
295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311
|||||
745 GCTGCAAAATATAGACCATATGGTGTTTTATAAGACTCCCGCAGCATCTCT 794

```

```

311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrp 328
|||||
795 TGTAAATACATAAATTTAGCACATCAGGGGTAGAGCCTGCAAGCACATATC 844
|||||
328 roAspLeuGlyLeuProProGluTyrGlyAlaLeuGluTyrValPhe 344
|||||
845 CTGACCTTGGGTGCCACCTGAATGATGAGCTCTGGAGTGGGTATTC 894
|||||
345 ProGluTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361
|||||
895 CCTGAATGGCGAGGAGCATGCCCTTGACAAGGGTGAGCAGCTTAATTT 944
|||||
361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnG 378
|||||
945 TTTGAAAGGTGCAGTTGTGACAGCAGATCGAATCGTGACTGTCAAGTAAG 994
|||||
378 lyTyrSerTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394
|||||
995 GTTATTCTGGGAGGTACAACTGCTGAAGGTGACAGGGCCTCAATGAG 1044
|||||
395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411
|||||
1045 CTCTTAAGCTCCAGAAAGAGTGTATTAAACGGAATTGTAATGGAATTGA 1094
|||||
411 pIleAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428
|||||
1095 CATTAATGATTGGAAACCTGCCACAGACAATGTATCCCTGTCAATATT 1144
|||||
428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
|||||
1145 CTGTTGATGACCTCTCTGGAAGGCCAATGTAAAGGTGCATTCGAGAAG 1194
|||||
445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG1 461
|||||
1195 GAGCTGGGTTAACCTATAAGGCTGATGTCTCTGATTGGCTTTATTGG 1244
|||||
461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478
|||||
1245 AAGTTGGATTATCAGAAAGGCATTCATCTCAACTTATCATACCGAG 1294
|||||
478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494
|||||
1295 ATCTCATCGGGAGAGATGTTCAATTTGTCATGCTTGGATCTGGTGACCCA 1344
|||||
495 IlePheGluGlyTyrMetArgSerThrGluSerSerTyrLysAspLysPh 511
|||||
1345 GAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAAT 1394
|||||
511 eArgGlyTyrPvalGlyPheSerValProValSerHisArgIleThrAlag 528
|||||
1395 TCGTGGATGGTTGGATTAGTTGTTCCAGCTTTCCACCACCAATCAATGCCG 1444
|||||
528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544
|||||
1445 GCTGCGATATATTGTTAATGCCATCCAGATTCGAACCTTGTGTGCTCAAT 1494
|||||
545 GlnLeuTyrAlaMetClnTyrGlyThrValProValValHisGlyThrG1 561
|||||
1495 CAGCTATATGCTATGAGTATGGCAGTCTCTGTTGCTCCATGCAACTGG 1544
|||||
561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578
|||||
1545 GGGCCTTAGAGATACCGTGGAGAACTTCAACCTTTCCGCTGAGAAATGGAG 1594
|||||
578 luGluGlyThrGlyTyrPalaPheSerProLeuThrValAspLysMet.le 594
|||||
1595 AGCAGGGTACAGGGGTGGGCTTCGCACCCCTAACCCACAGAAAACATGTTT 1644
|||||
594 utrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSert 611
|||||
1645 GTGGACATTTCGGAAGTGCNAATATATACATACAGGGAACACAAAGTCTCT 1694

```

611 rpgluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627
|||||
1695 GGAAGGCGTAATGAAGCGAGGATGTCAAAGACTTCACGTGGGACCAT 1744
628 Ala 628
|||
1745 GCC 1747

seq_name: gb_pat:A63308

seq_documentation_block: 2992 bp DNA PAT 12-MAR-1998
LOCUS A63308

DEFINITION Sequence 1 from Patent WO9720936.

ACCESSION A63308

VERSION A63308.1 GI:3717138

KEYWORDS

SOURCE

ORGANISM

unidentified.

unclassified.

REFERENCE 1 (bases 1 to 2992)

AUTHORS Keeling, P.L. and Knight, M.E.

TITLE MODIFICATION OF STARCH SYNTHESIS IN PLANTS

JOURNAL Patent: WO 9720936-A 1 12-JUN-1997;

ZENSCA LTD (GB)

COMMENT Other publication AU 1037197 19970627.

FEATURES

Source

1. .2992

/organism="unidentified"

/db_xref="taxon:32644"

/clone="NUMBER 1"

BASE COUNT 758 a 655 c 801 g 776 t 2 others

ORIGIN

alignment_scores:

Quality: 2655.50 Length: 745

Ratio: 4.304 Gaps: 11

Percent Similarity: 82.819 Percent Identity: 71.544

alignment_block:

US-09-674-824-2 x A63308 ..

Align seg 1/1 to: A63308 from: 1 to: 2992

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20

|||||

614 TCGCGCGTGGCGCGCGTGCCTC.....CTCCTCGCGCG 648

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAlaLeuA 37

|||||

649 GGNGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698

37 rgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPro 53

|||||

699 AGCGCGTCTCGCGCGCGCGCGTGCCTCGCGGAGCTGAGCAGGAGGCGCC 748

54 AlaAlaArgProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 70

|||||

749 GCGCGCGCGCGCGTGCACCGCGCGTGTGGCGCGCGCGCGCGCGCGCG 798

70 yPheLeuAlaProProProProProProProProProProProProPro 87

|||||

799 CTTCCTCGCG...CGCGCGCGCGAGCCACCGCGGTGAGCGCGCATCGACGC 845

87 InProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeu 103

|||||

846 CGCGCGCGCGTGCACCGCGCGCGTGGGGACCTCGGT.....CTCGAA 889

104 LeuGluGlyLeuAlaGluAspSerIleAspSerIleValAlaAlaSe 120

|||||

890 CCTGAAGGAGTGTGAAGTTCATCATCAACACAGTAGTGTGGCAAG 939

120 rGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAlaLys 137

|||||
940 TGAGCAAGATTCTGAGATTGTGGTGGAAAGGAGCAAGCTCGAGCTAAAG 989
137 alThrArgSerIleValPheValThrGlyGluAlaPro.TyrAlaLys 153
|||||
990 TAACACAAACATGCTTTGTAAACCGCGAAGCTTCTCCTTAATCGAAA 1039
153 sSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAla 170
|||||
1040 GTCTGGGGTCTAGGAGATGTTGTGGTTCATTGCCAGTGGCTCTTGTCTG 1089
170 laArgGlyHisArgValMetValValMetProArgTyrLeuAsnGlySer 186
|||||
1090 CTCGCGGTACCGGTGATGTTGTAATGCCAGACATTTAAATGGTATACC 1139
187 SerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLys 203
|||||
1140 TCCGATAAGAAATATGCAATGCATTTTACTCAGAAAAACACATTCGGAT 1189
203 eProCysPheGlyCysSerHisGluValThrPhePheHisGluTyrArg 220
|||||
1190 TCCATTCTTTGGCGGTGAACATGAAGTTACCTTTCTCCATGAGTATAGAG 1239
220 sPasnValAspTyrValPheValAspHisProSerTyrHisArgProGly 236
|||||
1240 ATTGAGTTGAGTGGTGTGTTGTTGATCATCCTCATATCACACACCTGGA 1289
237 SerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTy 253
|||||
1290 AATTATATGAGATAAGTTGGTGGCTTTTGGTGATTAATCAGTTTCAGATA 1339
253 rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeu 270
|||||
1340 CACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGGTCTTGAATGG 1389
270 lyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHis 286
|||||
1390 GAGGATATATTATGGACAGAAATTCATGTTGGTTGTCATGATGGCAT 1439
287 AlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrGlyVa 303
|||||
1440 GCCAGTCTAGAGCCAGTCTTCTTGTGCTGCAAAATATAGACCATATGGTGT 1489
303 lTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnG 320
|||||
1490 TTATAAAGACTCCCGCAGCATTTCTTTAATACATAAATTTAGCATCAGG 1539
320 lyValGluProAlaSerThrTyrProAspLeuGlyLeuProGluTrp 336
|||||
1540 GTGTAGAGCTGCAAGCACATATCTGCACCTTGGTGGTGGCCACCTGAATGG 1589
337 TyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLe 353
|||||
1590 TATGGAGCTCTGGAGTGGTATTCCTCAATGGCGGAGGAGGATGCCCT 1639
353 uAspLysGlyGluAlaValAsnPheLeuLysGlyValAlaValThrAla 370
|||||
1640 TGCAAGGGTGAGGAGTAAATTTTCAAGGTGTCAGTTGTGACAGCAG 1689
370 spArgIleValThrValSerGlnGlyTyrSerTrpGluValThrAla 386
|||||
1690 ATCCAAATCGTACGTGAGTAAGGTTTATTCATGGAGGTCCACAACGTCT 1739
387 GluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLe 403
|||||
1740 GAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATT 1789
403 uAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrAla 420
|||||
1790 AAACGGAATTTGAATGAATTTGACATTAATGTTGGAACCTGCCACAG 1839
420 spLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAla 436
|||||


```

1840 ACAATGTATCCCTGTCTATTCTGTGTGACCTCTCTTGAAGGCT 1889
437 LysCysLysAlaGluLeuGlnLysGluLeuProValArgGluAs 453
1890 AAATGTAAGGTGCATTTCAGAAGGAGGTGGGTTTACCTATAAGGCTGA 1939
453 pValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleA 470
1940 TGTTCCTCTGATTGGCTTTATTGGAAGATTGGATTATCAGAAAGGCATTG 1989
470 spLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPhe 486
1990 ATCTCATTCAACTATATACAGATCTCATCGGAAGATGTTCAATTT 2039
487 ValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerTh 503
2040 GTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTAC 2089
503 rGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValp 520
2090 AGAGTCGATCTTCAAGGATAAATTCGTGGATGGTTGGATTAGTCTTC 2139
520 roValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSer 536
2140 CAGTTTCCACCGAATACTCGGCTGGCGATATATTGTTAATGCCATCC 2189
537 ArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyTh 553
2190 AGATTGGAACCTTGTGCTCAATCAGCTATATGCTATGCAGTATGGCAC 2239
553 rValProValValHisGlyThrGlyLeuArgAspThrValGluThrP 570
2240 AGTTCTCTGTTCTCATGCACTGGGGGCTTAGAGATACCGTGGGAAC 2289
570 heAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSer 586
2290 TCACCCCTTCGGTCAGANTGGAGGAGGAGGTACAGGGTGGGCATTCGCA 2339
587 ProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSert 603
2340 CCCCTAACACAGAAAAATGTTGTGGACATTGCGAATTCGAATATCTA 2389
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
2390 CATACAGGGAACACAGTATAATGGGAGGGCTTAATGAAGCCAGGCATG 2439
620 ThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSerS 636
2440 TCAAAAGAGTTCACGTGGGACCATGGCGCTGAACAATACGACAAATCTT 2489
636 erSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG 652
2490 CCAGTGGGCTTCATCGGATCGACCC...GATGTTCAATGGAAAAAGG 2536
652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerp 669
2537 ACCAAAGTTGGTTGGTTCCTTGAAGATTATCAGTTCATCATCTATAGTA 2586
669 heArgGlyProGluGlyTyPro.....CystThrLeuArgCysProAla 683
2587 AGCTGAATGATGAAGAAAAACCCCTGTACATTACATGGAGGCAGACCGG 2636
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPhe..... 695
2637 CTATTGGCTCCATTGCTCAATGTCGCTTTGGCTTGCCTTCGCTCGATGG 2686
695 ..... 695
2687 ACCGGATGCAGTGAGGAATCCAGNCGACGACAGTTTGAAGGATAGGAA 2736
696 .....AlaGlySerArgThrTyrAspGlyCysAlaAlaValThr 710
2737 GGGGAGCTGGAAGC...AGTCACGCGAGGCAAGCCTTCGCCGTTAAT 2783

```

```

711 AlaSerGlyGlyArgGlnLeuGlnPhe 719
::: |||
2784 TCATATGGAACAAGCTGGAGTCAGTTT 2810

```


OM of: US-09-674-824-2 to: N_Geneseq_1101.* out_format : pfs
Date: Mar 28, 2002 5:41 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame2_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09674824/runat_27032002_094252_18454/app_query.fasta_1.824
-DB=N_Geneseq_1101 -QFMT=fastap -SURFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09674824_@CNI_1_261 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-674-824-2
Query length: 756
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 180.010000

seq_list:

Sequence	Strd Orig	Zscore	EScore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ24487			+ 4044.00	4329.98	4.2e-233
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ34651			+ 3930.00	4208.06	2.6e-226
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV01528			+ 3612.00	3868.22	2.2e-207
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50636			+ 2826.50	3023.93	2.3e-160
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ45183			+ 2821.50	3018.41	4.7e-160
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ50636			+ 2809.50	3006.08	2.3e-159
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50643			+ 2806.50	3004.41	2.9e-159
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV29756			+ 2709.50	2901.51	1.5e-153
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV70960			+ 2708.50	2900.46	1.8e-153
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV67285			+ 2655.50	2838.67	9.0e-150
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV68832			+ 2651.00	2833.85	9.0e-150
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV29759			+ 2346.50	2727.23	7.9e-144
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50637			+ 2495.00	2672.47	8.8e-141
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ34652			+ 2349.00	2498.33	4.4e-131
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50642			+ 2036.00	2180.38	2.3e-113
/SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT:AAZ46797			+ 1954.00	1981.90	2.6e-102
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV32325			+ 1731.50	1851.49	4.7e-95
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV70958			+ 1252.50	1334.31	3.0e-66
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50651			+ 1246.00	1328.64	6.3e-66
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50646			+ 1239.50	1321.03	1.7e-65
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1212.00	1292.48	8.5e-64
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1209.00	1290.32	6.5e-64
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1203.00	1279.41	3.5e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1203.00	1279.41	3.5e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1200.50	1277.14	4.6e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1200.50	1277.14	4.6e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1182.50	1257.75	5.6e-62
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1154.50	1230.36	1.9e-60
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1154.50	1230.36	1.9e-60
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1076.00	1146.53	8.7e-56
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV29754			+ 1076.00	1146.53	8.7e-56
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV7286			+ 1071.00	1140.81	1.8e-55
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV66833			+ 1071.00	1140.81	1.8e-55
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV66833			+ 943.50	1002.72	8.9e-48
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV67287			+ 943.50	1002.35	9.4e-48
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV29753			+ 900.00	955.42	3.9e-45
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV29753			+ 887.50	943.04	1.9e-44
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV51178			+ 881.00	921.39	3.0e-43
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ60319			+ 870.50	926.32	1.6e-43
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ32824			+ 863.00	917.37	5.1e-43

/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ96227 - 841.50 892.51 1.2e-41
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ96431 + 788.00 823.65 8.4e-38
/SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT:AAQ25978 + 678.50 722.63 3.6e-32
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV23619 + 678.50 722.63 3.6e-32
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV70951 + 634.50 675.24 1.6e-29
1488

seq_name: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ24487

seq_documentation_block:

ID AAZ24487 standard; DNA; 2805 BP.

XX AAZ24487;

AC XX

DT 18-FEB-2000 (first entry)

XX Wheat soluble starch synthase DNA.

XX Soluble; starch synthase; wheat; transgenic plant; starch production;
KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;
KW adhesive; textile; building material; soil stabilizer; wetting agent;
KW fertilizer; plant-protection; cosmetic; flocculant; ss.

XX Triticum aestivum.

XX Key Location/Qualifiers

FT CDS 314..2584

FT /*tag= a

FT /product= "soluble starch synthase"

XX DE19820607-A1.

XX 11-NOV-1999.

XX 08-MAY-1998; 98DE-1020607.

XX 08-MAY-1998; 98DE-1020607.

XX (AGRE) HOECHST-SCHERING AGREVO GMBH.

XX Loez H, Luetticke S, Block M;

XX WPI; 2000-024508/03.

XX P-PSDB; AAY50818.

XX New enzyme with starch synthase activity, useful for producing starch
for foods and packaging materials -

XX Claim 1b; Page 15-19; 24pp; German.

XX This invention describes a novel protein (I) with the activity of wheat
starch synthase. Transgenic plants, specifically wheat, that contain (I)
are used for production of starch, used particularly in foods,
particularly baked and pastry goods and for making packaging materials or
disposable items. Starch may also be used as starting materials for
glucose or glucan components (e.g. for fermentation or further chemical
conversion); in paper and pulp production, as adhesives, in textiles,
in preparation of gypsum-based building materials, as soil stabilizer,
as wetting agent etc. in fertilizer and plant-protection compositions,
as binder (in pharmaceuticals, cosmetics, coal briquetting and casting
sand), as flocculant in soil or coal slurries, as rubber and leather
additives, and for production of synthetic polymers, e.g. polyurethane
films. Transgenic plants with increased/decreased production of (I)
produce starches with altered physical and/or chemical properties such as
amylose/amylopectin ratios, degree of branching, mean chain length,
phosphate content, gelatinization properties, gel- or film-forming
properties, or starch grain size or structure. This sequence encodes the
soluble starch synthase isolated from wheat (Triticum aestivum L. cv.
Florida).

XX Sequence: 2805 BP; 683 A; 703 C; 763 G; 656 T; 0 other;

XX alignment_scores:

Quality: 4044.00 Length: 756
 Ratio: 5.349 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-674-824-2 x AA24487 ..

Align seg 1/1 to: AA24487 from: 1 to: 2805

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
 314 ATGGCGCGAGCGCGCTCGCGCGCGCGCTCGCGCGCGCGCGCGCGCT 363
 17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
 364 CGCGCGCGATCGCGCGAGCGCGCGCGCGCGCGCGCTCGCGCTCGCG 413
 34 laArgLeuArgAlaAlaArgGlyArgTyrValAlaGluLeuSerArg 50
 414 CGCGCGCTCGCGCGCTTGGCGCGCGCGCTACGTGCGCGAGCTCAGCAGG 463
 51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnGlnGlnGlnGln 67
 464 GAGGCT 513
 67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
 514 CGTGCGCGAGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 563
 84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
 564 CCCGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 613
 101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
 614 GACCTCCCTGCTGGAAGGATTGCTGAGATTCCATCGACACATAATTGT 663
 117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProg 134
 664 GCCTCAAGTAGCAGGATTCTGAGATCATGGATGCGGAATGAGCAACCTC 713
 134 InAlaIlyValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
 714 AAGCTAAAGTTACAGTACGATGCTGTTGTTGACGTTGTTGAGCTGCTCCT 763
 151 TyrAlaIysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
 764 TATGCAAACTCAGGGGGTTGGGAGATGTTGTGCTTACCAATTCG 813
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
 814 TCTTGCTGCTGCTGCTACCGAGTATGTTGTAATGCCAAGATCTTAA 863
 184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
 864 ATGGGTCCTCTCATATAAACTATGCCAAGGCATATACACTCGGAAGCAC 913
 201 ileLysileProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
 914 ATTAAGATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATCA 963
 217 uTyrArgAspAsnValAspTyrValPheValAspHisProSerTyrHisA 234
 964 GTATAGAGACACGTCGATTGGGTGTTTGTGATCATCCGTCATATACACA 1013
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
 1014 GACCAGGAAGTTATATGAGAGATAAATTTGGTGCTTTTGTGATATAATCAG 1063
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
 1064 TTCAGATACACACTCCTTTTGTCTATGTGCTGATGCGAGGCGGCGGCTT 1113

267 uGluLeuGlyGlyTyrIleTyrGlyClnAsnCysMetPheValValAsnA 284
 1114 TGAATTGGGAGGATATATTTATGGCAGAAATTCATTTTGTGTGAAGC 1163
 284 sPTTrPHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
 1164 ATTGGCATGCCAGCCTGTGCCAGTCTCTTCTGCTGCAAAATATAGACCA 1213
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
 1214 TAGCGTGTTCACAGACATCCCGCAGCAGCTTGTATACATAAATTAGC 1263
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 334
 1264 ACATCAGGGTGTGGAGCCTGCAAGTACATATCTGATCTGGGATTCGCTC 1313
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
 1314 CTGAATGGTATGAGCTTTAGAAATGGTATTTCCACAATGGGCAAGGAGS 1363
 351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
 1364 CATGCCCTTCACAAGGTGAGGAGTTAACTTTTGAAGGAGCAGTGTGT 1413
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
 1414 GACAGCAGATCGGATTGTGACCGTCACTCAGGCTTATTCATGGGAGGTCA 1463
 384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuSerSerArgLys 400
 1464 CAATGCTGAAGGTGGACAGGCTCAATCAGCTCTTAAGCTCCGCAAAA 1513
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
 1514 AGTGATTTGAATGGAATGTAAATGGAATGACATTAATGATTGGAAACCC 1563
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
 1564 CACCACAGACAAAGTGTCTCCTCATCATATTATCTGTCGATGACCTCTCTG 1613
 434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
 1614 GAAAGGCCAAATGTAAAGCTGAATTCAGAAAGGAGTTGGGTTTACCTGTA 1663
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
 1664 AGGAGGAGATGTCCTCTCATTTGGCTTTATTCGAAGACTGGATTAACAGAA 1713
 467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
 1714 AGGCATTGATCTCATTTAAATGGCCATTCACAGAGCTCATAGGAGGAGCG 1763
 484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
 1764 TGCAATTTGTTCATGCTTGGATCTGGGATCCCAATTTTGAAGGCTGGATG 1813
 501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
 1814 AGATCTACCGAGTCGAGTTACAAGGATAAATTCGCTGGATGGTGGGATT 1863
 517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
 1864 TAGTCTCCAGTTTCCACAGAAATAACTGCAGGTTGCGATATATTTGTTAA 1913
 534 eProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
 1914 TGCCATCCGAGATTGAACCTTCGCGCTCTTAATCAGCTATATGCTATGCAA 1963
 551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
 1964 TATGTTACAGTTCTCTGTAGTTTCATGGAACCTGGGGGCTCCGAGACACAGT 2013
 567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584

```
|||||
2014 CGAGACCTTCAACCTTTGTTGTCAAAAGGAGGAGGATACAGGTGG 2063
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
2064 CGTTCTCACCGCTAACCGTGACAAAGATGTTGTGGCATTCGGAACCGG 2113
601 MetSerThrPheArgLysHisLysProSerTrpGluGlyLeuMetLysAr 617
2114 ATGTCGACATTGAGGAGCACAAGCGTCTGGGAGGGCTCATGAAGCG 2163
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSera 634
2164 AGGCATGACGAAGAACCATACGTGGGACCATGCCCGGAGCATGACGACA 2213
634 rgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650
2214 GATCTTCAGTGGGCTTCGTGGACCAACCTACGTATGATGAGCGGGA 2263
651 LeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSe 667
2264 CTGGGAGGTCAAAGTGGAGTCTCTTCAGCTCTGAAGACATCTCTTC 2313
667 rSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAlar 684
2314 ATCTTCCGCGCGCCGGAAGGATACCCCTGTACATTGCGTTGCTGCTA 2363
684 hrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThr 700
2364 CAGTAGAGTCCGCAATGCGCTGCTGTTGGTTGCGCGGTTCGAGAACA 2413
701 TyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnLe 717
2414 TATGACGGCTGTGCTGTCGGCGGTGACAGCTTCGGGTGACGACAGTT 2463
717 uGlnPheTrpGlyTleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaL 734
2464 ACAGTTTGGGAATAAGGAGGATGCTGTCAGGATGTTAACAACAA 2513
734 ysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleArg 750
2514 AGCACCACTCAGATGGCAGCTCTCTGTCGTGTACAGCTGAATCAGA 2563
751 AsnGlnLeuValThrIeu 756
2564 AACCACTGGTGACTCTT 2581
```

seq_name: /STD51/gcgdata/geneseg/geneseqn/NA1999.DAT:AA34651

seq_documentation_block:

ID AAX34651 standard; cDNA; 2662 BP.

XX AC AAX34651;

XX DT 05-JUL-1999 (first entry)

XX DE cDNA sequence of wheat starch soluble synthase I (SSS I).

XX KW Starch biosynthetic pathway; cereal plant; enzyme; SBE; DBE; GBSS;
KW starch branching enzyme; starch soluble synthase; debranching enzyme;
KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
KW grain softness protein I; bacterial isoamylase; glycogen synthase;
KW WSBE I-D4 gene; ss.

XX OS Triticum tauschii.

XX PN WO9914314-Al.

XX PD 25-MAR-1999.

XX PF 11-SEP-1998; 98WO-AU00743.

XX PR 20-MAR-1998; 98AU-0002509.

PR 12-SEP-1997; 97AU-0009108.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
PA (AUSU) UNIV AUSTRALIAN NAT.

XX Li Z, Morell M, Rahman S;
XX WPI; 1999-229525/19.

XX New isolated cereal plant enzyme genes used for, e.g. expression of
PT antisense sequences of granule bound synthase
XX

PS Claim 10; Page 81-83; 171pp; English.

CC The invention relates to a novel enzyme of starch biosynthetic pathway
CC in a cereal plant, where the enzyme is selected from starch branching
CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching
CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
CC SBE I of rice or maize. The methods and products can be used for
CC targeting expression specifically to the endosperm of the seeds of cereal
CC plants such as wheat or barley. They can be used for the expression of
CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
CC can be used for modifying the characteristics of starch produced by a
CC plant. The present sequence represents the wheat SSS I cDNA sequence.

XX Sequence 2662 BP; 592 A; 677 C; 754 G; 639 T; 0 other;

alignment_scores:

Quality: 3930.00 Length: 757
Ratio: 5.247 Gaps: 1
Percent Similarity: 98.943 Percent Identity: 98.415

alignment_block:

US-09-674-824-2 x AAX34651 ..

Align seg 1/1 to: AAX34651 from: 1 to: 2662

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgle 17
|||||
257 ATGGCGCGCAGCGGCTCGCGCGCGGTCCTCGCCCGCCAGCGTCGCGCT 306

17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
307 GCGCGCGGATCCGCGACGCGCGCGCGCGCTCGCGCTCGCTCGCGCGG 356

34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
357 CGCGGCTCGCGGCTTGGCGCGCGCGCGCTACGTTGCGGAGCTCAGCAGG 406

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProle 67
|||||
407 GAGGCT 456

67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
|||||
457 CGTGCGAGGCTTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506

84 laProThrGlnProProLeuProAlaGlyValGlyGluLeuAlaPro 100
|||||
507 CCG 556

101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
|||||
557 GACCTCTGCTCGAAGGAGTTCCTGAGGATTCATCGACACATAATTGT 606

117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnPro 134
|||||
607 GGCTGCAAGTGAAGGATTCATGATCATGATGATGATGATGATGATGATG 656

134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
 657 AAGCTAAAGTTACACGTAGCATCTGTTTGTGACTGGTGAAGCTGCTCT 706
 151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
 707 TATGCAAACTCAGGGGGCTGGGAGATGTTGTGCTGTTACCAATTGC 756
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
 757 TCTTGCTCTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
 184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
 807 ATGGGTCTCTCGATAAAACATATGCAAGGCAATATATACCTGGAGAC 856
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
 857 ATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTTTCATGA 906
 217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
 907 GTATAGACACAACGTCGATGGTGTGCTGATCATCCGTCATATCATA 956
 234 tGProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
 957 GACCAAGGAAGTTATATGAGAGATAATTTTGGTGTGCTTTTGGTGATAATCAG 1006
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
 1007 TTCAGATACACACTCTCTTTGCTATGCTGCTGCTGCTGCTGCTGCTGCT 1056
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
 1057 TGAATTGGGAGGATATTTATGGAAGAGATGTCATGTTGTGTGGAAGC 1106
 284 sPTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
 1107 ATTGGCATCCAGCTTGTGCGAGTCTCTTCTGCTGCAAAATATAGACCA 1156
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
 1157 TACGGTGTTTACAGAGATTTCCCGCAGCACCTTGTGTATACATAATTTAGC 1206
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334
 1207 ACATCAGGCTCTGGAGCCTGCAAGTACATATCTGATCTGGGATTGCCAC 1256
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
 1257 CTGAATGGTATGGAGCTTTAGAAATGGGTATTTCCAGAAATGGCAAGGAGG 1306
 351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyValaValVa 367
 1307 CATGCCCTTGCAGAGGCTGAGCAGTTACTTTTGAAGGAGCAGCTCGT 1356
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
 1357 GACAGCAGATCGAATTTGTGACCTCAGTCAGGCTTATTCAGGAGGTCA 1406
 384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
 1407 CAACCTGCTAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAANA 1456
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
 1457 AGTGTATTTGAATGGAATTTGAATGGAATTTGAATGGAATTTGAATGGA 1506
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
 1507 CACCACAGACAAGTGTCTCCCTCATCATTTCTGTCGATGACCTCTCTG 1556

434 lYlAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
 1557 GAAAGCCAAATCTAAAGCTGAATTCAGAAAGGAGCTGGGTTTACCTGTA 1606
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
 1607 AGGAGGAGATGTTCTCTGATTGGCTTTATTTGGAAGACTGGATTACCAGAA 1656
 467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
 1657 AGCATTTGATCTCATTAANATGCCATTCCAGAGCTCATGAGGAGGACG 1706
 484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
 1707 TGCAGTTTGTCTGATCTGGGATCTGGGATCCAAATTTTGAAGCTGGATG 1756
 501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
 1757 AGATCTACCGAGTCGAGTTACAAGGATAAATTCCTGGATGGGTGGATT 1806
 517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
 1807 TAGTGTTCAGTTTCCACAGATAACTGCAGGTTGCGATATATTTGTTAA 1856
 534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
 1857 TGCCATCCAGTTTGAACCTTGTGCTTAAATCAGCTATATGCTATGCCAA 1906
 551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567
 1907 TATGGTACAGTTCTGTAGTTTCATGGAACCTGGGGCTCCGAGACACAGT 1956
 567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
 1957 CGAGACCTTCAACCTTTTGTGTCAAAAGGAGAGAGGAGGTACAGGGTGG 2006
 584 lPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
 2007 CGTTCTACCCCTTACCGTGGACAAAGATGTTGTGGCATTCGGAACCGCG 2056
 601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
 2057 ATGTCGACATTCAGGAGACACAAGCCGCTCTGGAGGGGCTCATGAAGCG 2106
 617 gGlyMetThrLysAspHisThrTrpAspHisAlaA.ProSerSerThrSer 633
 2107 AGGCATGACGAAGAACCATACGTGGACCATGCCGCCGAGCAGTACGAGC 2156
 634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgL 650
 2157 AGATCTCGAATGGCCCTTCGTGGACCAACCTACGTCTAGACGGGG 2206
 650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
 2207 ACTGGGAGGTGCAAGCGCGGCTCTCTTGGAGCTCTGAAGACATGTTCT 2256
 667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683
 2257 CATCTTCCGGCGCCCGAAGGATACCCCTGTACATTGCGTTGCTCTGCT 2306
 684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700
 2307 ACAGTAGAGTCGAATGCGCTGCTTGTCT.TGGTCCCGCGGTTCGAGAGT 2355
 700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717
 2356 ACATGACGGCTGTGCTGCTGCGCGGTGACAGCTTCGGGTGGATGACAGT 2405
 717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733
 2406 TACAGTTTTGGGGAATAGGAAGGATGTGCTGCGAGGATGGTTAACAGCA 2455
 734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750

|||||
2456 AAGCACCACTAGATGGCAGCCTCTCTCGTGTATACAGTGAATCAG 2505
750 gAsnGlnLeuValThrLeu 756
|||||
2506 AAACCACTGGTACTCTT 2524

seq_name: /STD51/gcgdata/geneseq/geneseq/NA1998.DAT:AAV01527

seq_documentation_block:

ID AAV01527 standard; cDNA to mRNA; 2239 BP.

AC AAV01527;

DT 21-MAY-1998 (first entry)

DE Wheat soluble starch synthase partial cDNA sequence.

KW Starch synthase; wheat; transgenic plant; ss.

OS Triticum aestivum L. cv. Florida.

FH Key Location/Qualifiers

FT CDS 3..2018

FT /*tag= a

PN W09745545-Al.

XX 04-DEC-1997.

XX 28-MAY-1997; 97WO-EP02793.

XX 11-SEP-1996; 96DE-1036917.

PR 29-MAY-1996; 96DE-1021588.

PA (AGRE) HOECHST-SCHERING AGREVO GMBH.

XX Block M, Loerz H, Luetticke S, Froberg C, Kossmann J;

PI Walter L;

XX WPI; 1998-032652/03.

DR P-PSDB; AAW23937.

PT Nucleic acid encoding starch synthase enzymes from wheat - for
PT transgenic plants that produce modified forms of starch, useful e.g.
PT in foods, or for production of packaging materials and disposable
PT goods

PS Claim 1; Page 47-51; 71pp; English.

XX This near full-length cDNA clone, designated TaSSS, codes for a
CC soluble starch synthase (see AAW23837) of summer wheat (cv. Florida).
CC It was isolated from a phage cDNA library of 21-day-old wheat
CC caryopses by screening with a PCR fragment derived from rice soluble
CC starch synthase (see also AAV01529-30). A second clone (see AAV01528),
CC coding for wheat granule-bound starch synthase (see AAW23938) is also
CC claimed. These isolated nucleic acids can be inserted into vectors
CC for production of transgenic plants, particularly starch-producing
CC plants, specifically wheat. Use of the isolated nucleic acids, or
CC of antisense sequences, allows starch metabolism to be regulated in
CC transgenic plants. Overexpression may result in improved crop
CC yield, while modification of starch in planta may eliminate the
CC need for subsequent chemical/physical modification. Plants with
CC altered levels of the various isoforms of starch synthase will
CC produce starch of different chain length, amylose/amylopectin ratio,
CC degree of branching, phosphate content, gelatinisation behaviour,
CC granule size and shape, viscosity etc. The starch produced by such
CC plants is useful particularly in foods or to produce packaging
CC materials or disposable goods, as well as in any other known use of
CC starch.

XX Sequence 2239 BP; 611 A; 448 C; 590 G; 590 T; 0 other;

alignment_scores:

Quality: 3612.00 Length: 671

Ratio: 5.383 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-674-824-2 x AAV01527 ..

Align seg 1/1 to: AAV01527 from: 1 to: 2239

86 ThrGlnProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102
|||||
3 ACGCAGCGCCCTCGCGGACCGCGCGTGGGGAACTCGCGCCGACCT 52
102 uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAla 119
|||||
53 CCTGCTCGAAGGGATTGCTGAGGATTCATCGACAGCATAAATTGGCTG 102
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAla 135
|||||
103 CAAGTGAGCAGGATTCTGAGATCATGGATGCGAATGAGCAACCTCAAGCT 152
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152
|||||
153 AAAGTTACACGTAGCATCGTGTGTGACTGGTGAAGCTGCTCCTTATGC 202
152 aLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeu 169
|||||
203 AAAGTCAGGGGGTTGGGAGATGTTGTGTTGCTTACCAATTGCTCTTG 252
169 laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185
|||||
253 CTGCTCGTGTACCAGAGTGTGTTGTAATGCCAAGATACTTAATGGG 302
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysIleLeLy 202
|||||
303 TCCTCTGATAAAACATATGCAAGGCAATTATACACTCGGAAGCACATTA 352
202 sileProCysPheGlySerHisGluValThrPhePheHisGluTyrA 219
|||||
353 GATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGAGTATA 402
219 rgAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 235
|||||
403 GAGACAACGTCGATTGGGTCTTTGTCATCATCCGTCATATCACAGACCA 452
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252
|||||
453 GGAAGTTTATATGGAGATAATTTTGGTGTCTTTTGGTGATAATCAGTTCAG 502
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269
|||||
503 ATACACACTCCTTTGCTATGCTCATCGAGGCCCACTAATCCTTGAAT 552
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrp 285
|||||
553 TGGGAGGATATATTATGGACAGAAATGTCATGTTTGTGTGGAACGATTGG 602
286 HisAlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrG1 302
|||||
603 CATGCCAGCGCTTGTGCCAGTCTCTTGTGTCAAAATATAGACCATACGG 652
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
|||||
653 TGTTTACAGAGATTCGCCGAGCACCTCTGTATACATAATTTAGCACATC 702
319 InGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProGlu 335
|||||
703 AGGGTGTGGAGCGCTCAAGTACATATCTGATCTGGGATTGCTCTCTGAA 752
336 TrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgHisAl 352
|||||

753 TGGTATGGAGCTTTAGAAATGGGTATTTCCAGAAATGGGCAAGGAGGCATGC 802
352 aLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrA 369
803 CCTTGACAAGGGTCAGGACGTTAACTTTTGAAGAGGACGAGTTGTGACAG 852
369 laAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThr 385
853 CAGATCGGATTTGACCGCTCAGTCAGGCTTATTCATGGAGGTCACAACT 902
386 AlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402
903 GCTGAAGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCCGAAAAGTGT 952
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 419
953 ATTGAATGGAATTTGAAATGAATTTGACATTAATGATTGGAACCCACCA 1002
419 hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435
1003 CAGACAAGTGTCTCCCTCATCTATTCTGTGCGATGACCTCTCTGGAAG 1052
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgG1 452
1053 CCCAAATGAAGCTGAATTCAGAAAGAGTTGGTTTACCTGTAAAGGA 1102
452 uAspValProLeuLeuGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469
1103 GGATGTTCTCTGATTTGGCTTTATTGGAAGACTGGATTACCAGAAAGCA 1152
469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485
1153 TTGATCTCATTAATAATGCCATTCAGAGCTCATGAGGAGGAGCGTGCAA 1202
486 PheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgse 502
1203 TTTGTCATGCTTGGATCTGGGGATCCAAATTTTGAAGCTGGATGAGATC 1252
502 rThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerV 519
1253 TACCGAGTCGAGTTACAAAGATAAATTCGTTGGATGGGTTGATTTAGTG 1302
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535
1303 TTCCAGTTTCCCAAGATAACTGCGAGTTGCGATATATTGTTAATGCCA 1352
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG1 552
1353 TCGAGATTGAACCTTGGGCTCTTAATCAGCTATATGCTATGCAATATGG 1402
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569
1403 TACAGTTCTGTAGTTTCAATGGAATCGGGGCTCCGAGACACAGTCGAGA 1452
569 hrPreAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPhe 585
1453 CCTTCAACCCCTTTGGTGCACAAAGAGAGGAGGTTACAGGGTGGGGCTTC 1502
586 SerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSe 602
1503 TCACCGCTTAACCGTGGACAAAGATGTTGGGCAATGCGAACCGCATGTC 1552
602 rThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyM 619
1553 GACATTCAGGGAGCAACAGCGCTCTGGGAGGGGCTCATGAAGCGAGCA 1602
619 etThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSer 635
1603 TCACGAAAGACCATACGTTGGGACCATGCCGAGCAGTACGAGCAGATCT 1652
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG1 652
1653 TCGAGTGGGCTCTCTGGACCAACCCCTACGTCATGTAGACGGGACTGGG 1702

652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerP 669
1703 GAGTCCCAAGTGGAGTCTCTCTCAGCTCTGAAGACATCCTCTTCATCCT 1752
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685
1753 TCCGCGCGCCGGAAGGATACCCCTGTACATTCGCTTCTCTCTACAGTA 1802
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702
1803 GAGTCGCAATCGCCTGCTTGTGTTGTCGCCGTTTCGAGAACATATGA 1852
702 pGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnP 719
1853 CGGCTGTGCTGCTCGCGCGGTGACAGCTTCGGGTGGACGAGTTACAGT 1902
719 heTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
1903 TTTGGGAATTAAGAGGGATGCTGCTGAGGATGGTTAACAGCAAGACAC 1952
736 HisSerAspGlySerLeuSerValArgValThrAlaGluIleArgAsnG1 752
1953 CACTCAGATGCGACGCTCTCTGTCCGTGTACAGCTGAAATCAGAAACCA 2002
752 nLeuValThrLeu 756
2003 ACTGGTGACTCTT 2015

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AA250636

seq_documentation_block:

ID AA250636 standard; cDNA; 2491 BP.

XX AA250636;

XX 23-MAY-2000 (first entry)

DE Corn soluble starch synthase composite gene sequence.

XX Soluble starch synthase; starch fine structure; corn;

KW transgenic plant; amylose; amylopectin; amylose polymerisation;

KW non-granule bound starch synthase; non-GBSSI; altered starch; food;

KW paper; plastic; adhesive; ss.

XX Zea mays.

OS WO200006755-A2.

XX 10-FEB-2000.

XX 26-JUL-1999; 99WO-US16296.

XX 28-JUL-1998; 98US-0094436.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

PA Broglie KE, Lightner JE;

XX WPI; 2000-195311/17.

XX Producing transgenic cereal crops with altered starch structure useful

PT for preparing foodstuff, paper, plastic or adhesives, comprises

PT transforming crops with chimeric sense or antisense gene construct

PT encoding starch synthase

XX Claim 5; Page 51; 56pp; English.

XX The present sequence is the corn soluble starch synthase (SSI) composite
CC gene. This was used in the construction of plasmid pSS1 for the
CC generation of an antisense construct for suppression of SSI expression
CC in corn. The starch fine structure derived from a grain of the cereal
CC crop can be altered in the transformed cereal crop by changes in amylose


```
1674 TAGTGTTCAGCTTTCCACCGCAATAACTGCGGCTGCGATATATTGTAA 1723
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||
1724 TGCCATCCAGATTCCAACTTGTGCTTCAATACGTATATGCTATGCAG 1773
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||
1774 TATGCCACAGTCTCTGTGTCATCACTGCGGCTTAGAGATACCGT 1823
567 lgluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrTrpA 584
|||||
1824 GGAGAACTCAACCTTTCCGCTGAGATGGAGAGCAGGCTACAGGTTGG 1873
584 laphSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||
1874 CATTCGCACCCCTAACCCACAGAAACATGTTGTGACATTCGGAACGTGA 1923
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||
1924 ATATCTACATACAGGAACACAGTCTCTGCGGAGGCTAATCAAGCG 1973
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
|||||
1974 AGGCATGTCAAAAGACTTCACGCTGGGACCATGCCGCTGAACAATACGAAC 2023
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgL 650
|||||
2024 AAATCTTCAGTGGGCCCTTCATCGATCCGTCGTCGTCATGTAATAAAA.. 2071
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSers 667
|||||
2072 .....AGGACCAAGTGGTGGTTCCT.....TGAAGAT 2099
667 erSerPheArgGlyPro.....GluGlyTyrProCys 677
|||||
2100 CATCATGTCATCATCTCTAGTAACTAGTAAATGATGAAAGAAACCCCTGT 2149
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694
|||||
2150 ACATTACAT..... 2158
694 pPheAlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrA 711
2158 ..... 2158
711 laSerGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727
2159 .....GGAAGGAGACCGGCTATTGGCTCCATTGCTCCAACTCTGC 2200
728 AlaGlyTrpLeuThrAlaLysHis 735
2201 TTTGGCTGGCTTGCCTCGATGCAC 2224
```

seq_name: /SIDSl/gcgdata/geneseq/geneseq/NAL1994.DAT:AAQ45183

seq_documentation_block:

ID AAQ45183 standard; cDNA to mRNA; 2533 BP.

XX AC AAQ45183;

XX DT 16-NOV-1994 (first entry)

XX DE Soluble rice starch synthetic enzyme.

XX KW Rice; starch synthetic enzyme; transit peptide; amyloplast;

XX KW transition; proloplast; expression; ss.

XX OS Oryza sativa.

XX FT key

XX FT 5'UTR

XX FT 1..113

XX FT /*tag= a

```
FT CDS 114..1994
FT sig_peptide /*tag= b
FT 114..452
FT /*tag= c
FT mat_peptide 453..1991
FT /*tag= d
FT 3'UTR 1992..2533
FT /*tag= e
PN JP06070779-A.
XX 15-MAR-1994.
XX 07-JUL-1992; 92JP-0179947.
XX 07-JUL-1992; 92JP-0179947.
XX (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX WPI; 1994-128678/16.
XX P-PSDB; AAR51231.
XX Soluble rice starch synthetic enzyme gene and transit peptide -
XX for the efficient transport of heterologous proteins to
XX amyloplast
XX Claim 1; Page 15-17; 18pp; Japanese.
XX The soluble rice starch synthetic enzyme gene has, at the N-terminal,
XX the transit peptide that is required for the transition of this
XX enzyme to the amyloplast. Introduction of this gene into the rice
XX proloplast augments the expression of soluble rice starch synthetic
XX enzyme. The transit peptide coding sequence can be used for
XX the efficient transition of any protein into amyloplasts.
XX Sequence 2533 BP; 630 A; 550 C; 701 G; 652 T; 0 other;
```

alignment_scores:

Quality: 2821.50 Length: 726

Ratio: 4.522 Gaps: 18

Percent Similarity: 85.950 Percent Identity: 76.309

alignment_block:

US-09-674-824-2 x AAQ45183 ..

Align seg 1/1 to: AAQ45183 from: 1 to: 2533

2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17

|||||

126 GCGGGGATGGGATCGGGCGCGCTGTGGTGGCGCGCGAGGTGAGG... 173

17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34

|||||

174CCGGGGAGG.....AGGT 186

34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50

|||||

187 TCGCGCTCAGCGGTGGGAGG...CGGTGCGTGGCGGAGGTGAGCAGG 233

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65

|||||

234 GACGGTGGGTGGCG.....CAGGGCGGCTGGCACCGCGCGCC 271

66ProLeuValProGlyPheLeuAlaPro.....ProP 76

|||||

272 GCTGTGAAGCAGCGGCTCTCCGACCTTCTCGTGGCGAGGTGAGCGC 321

76 roProAlaProAlaGlnSer.....ProAlaProThrClnProLeu 90

|||||

322 CACCGCGCGCCAGCAGTGGCGCGCGCGCGCGCGCGCGCGCGGTG 371

91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyI 107

372 CGGACTCCGGCGTGGGAGATGAGCCCGAT.....CTAGAAGTGCT 415
107 eAlaGluAspSerIleAaspSerIleIleValAlaAlaSerGluGlnAsp 124
416 CACAGAAGATCCATCGACAAAAACANATTTTGTGGCTAGTAGCAGCAGGAGT 465
124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSer 140
466 CTGAGATCATGTGATGAAGGAGCAAGCTCAAGCTAAAGTAACACGCAGC 515
141 lleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157
516 GTTGCTTTGTAAACGGGTGAAGCTTCTCCTTATGCAAGTCAAGTGGACT 565
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisA 174
566 AGGAGATGTTGTGTGTTCCACTGCCAATTCCTTCTGCTCTCGTGGTCATC 615
174 rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190
616 GTGTGATGGTTGTAATGCCGAGATACATGAACGGGGCTTGAACAAAAAT 665
191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207
666 TTTGCCAAGCGCATTTTACACTGAAGCAGCATTAAGGATTCATGCTTTGG 715
207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspT 224
716 CGGACAACATGAAGTTACTTTTTCAGGAGTATAGGATTCGTCTGATT 765
224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
766 GGGTCTTTGTGATCATCCCTCATATCATAGACCTGGAATTTGATGGA 815
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
816 GATAATTTTGGTCTTTGGCGATAATCAGTTCAGATACACACTCCTGTG 865
257 stYrAlaLysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleT 274
866 CTATCGCGCGTGTGAAGCCCATTAATCTTGAACTGGGAGGATATATCT 915
274 yrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
916 ATGGACAGAAATGCATGTTCTGTGTAATGATTGGCATGGCAGTCTTG 965
291 ProValLeuLeuAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
966 CCAGTCTCTTGTGTCGCAAAATATAGACCATATGGTGTTTACAGGGA 1015
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
1016 CGCAGTCTTCTTGTCATACATAATCTAGCACATCAGGCTGGAGCCTG 1065
324 laSerThrTyrProAspLeuGlyLeuProProGluTrpTyrGlyAlaLeu 340
1066 CCAGTACATATCCTGACCTGGGATGGCACCTGAATGGTATGGAGCATTA 1115
341 GluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLysGlyG 357
1116 GAATGGGTGTTTCAGAGTGGGCAAGCGGCGATGCCCTTCACAAGGGTGA 1165
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
1166 GGCAGTCAATTTTTAAAGCGCGCAGTTGTGACACAGCAGATCGAATTTG 1215
374 hrValSerGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGlyGln 390
1216 CTGTGAGCCAGGGGTATTTCATGGGAGGTCACACTGCTGAAGGTGGGCAA 1265
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleVal 407

1266 GGCCCTCAATGAGCTCTTAAGCTCCCGGAAGAGTGTATTGAATGGAATTGT 1311

407 lAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424

1316 AATGTGAATTGACATTATGATTGGAACCCCATCCACAGACAAGTTCTCC 1365

424 rOHishistyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440

1366 CTTATCATTTCTCTGTATGACCTGTCCGGAAGGCCAAGTGTAAAGCT 1415

441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuII 457

1416 GAATTGCAGAAGAGCTGGGTTTACCTATAAAGGCCCATGTGCTCTGAT 1465

457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474

1466 TGGCTTTATTGGAAGATTGGACTATCAAAAAGGCATTGATCTAATTAAC 1515

474 eTAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490

1516 TTGCCATTCCAGATCTCATGGGGACAATATCAATTTCGTATCGTCTTGA 1565

491 SerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSery 507

1566 TCTGGTGACCCAGGTTTTTGAAGATGGATGAGATCCACAGAATCAGGTA 1615

507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524

1616 CAGGGATAAATTCGTGGATGGGTGGATTAGTTGTTCCAGTTTCCACC 1665

524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540

1666 GAATAACTGCAGGTTCGATATATTGTGTGATGCCATCCAGATTCCGAAC 1715

541 CysGlyLeuAsnGlnLeuTyrAlaMetClnTyrGlyThrValProValVa 557

1716 TGTGGCCCAATCAGCTATATGCTATGCAATATGGTACAGTGGCTGTGT 1765

557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574

1766 TCATGGAATCGAGGCCCTCAGAGATACAGTGGAGAAATTTAAACCGTTG 1815

574 lYAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590

1816 CTCGAAAGAGAGACAGGAGTACAGGTGGGCATTCTCGCCACTAACCAT 1865

591 AspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607

1866 GAAAAAATCGTGGGCATTTCGGATGGCAATTTTCACATACAGGGAAC 1915

607 lSLeuProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623

1916 ACAAGTCCTCTTGGGAGGCTCTAATGAAGCGAGGCATGTCAAGCGACT 1965

624 ThrTrpAspHisAlaProSerSerThrSerArgSerSerSerGlyProS 640

1966 ACATGGGACCATCGCCCTCACAGTATGAACAGCATCTTCGAATGGCGCT 2015

640 eTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLys... 655

2016 CATGATCAACCATATGTCATGTAATGGATTGGAAGAGACGCAAAAT 2065

656 CysGluSerPro.....SerAlaLeuLysThrS 665

2066 TCTCGGAGGACCTCAATCTTCCTGTCTTTCATGAGCGGAATGAACA 2115

665 eSerSerSerPheArgGly.....Pr 672

2116 TGTACACTACATGAAAGGAACCCAGTTATGCAAGGTTCACAAAGCATC 2165

672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689

2166 TCAAGGTTACCTTTAGTGG.....CTGTGACTTGTGCCA..... 2199

689 ysalacysLeuLeuTrpPhe 695
 2200ATATGGTTC 2208

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT: AAT95785

seq_documentation_block:

ID AAT95785 standard; cDNA to mRNA; 2383 BP.

AC AAT95785;

DT 22-MAY-1998 (first entry)

DE Maize starch synthase type I cDNA.

KW Maize; starch synthase type I; starch; ds.

OS Zea mays.

Key Location/Qualifiers

CDS 2..1951

FT /tag- a

FT /product= starch_synthase_type_I

XX DE19619918-A1.

XX 20-NOV-1997.

XX PF 17-MAY-1996; 96DE-1019918.

XX PR 17-MAY-1996; 96DE-1019918.

XX PA (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.

XX PI Froberg C, Kossmann J;

XX WPI: 1998-000821/01.

DR P-PSDB; AAW38218.

XX PT DNA encoding maize starch synthase type I protein - for producing transgenic plants

XX PS Claim 1: Pages 16-20; 23pp; German.

CC The present sequence encodes maize starch synthase type I, useful in the production of starch. Starch can be used in various conventional starch applications, e.g. starch hydrolysate products, foods, papermaking, adhesives, textiles, building materials, soil stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal briquettes, ore and coal slurries, foundry casting, rubber, leather and synthetic polymers. The enzyme produces a starch stated to have different physicochemical properties, especially viscosity and gelling properties, from wild type starch.

XX SQ Sequence 2383 BP; 582 A; 546 C; 633 G; 622 T; 0 other;

alignment_scores:

Quality: 2809.50 Length: 749
 Ratio: 4.411 Gaps: 10
 Percent Similarity: 85.047 Percent Identity: 73.565

alignment_block:

US-09-674-824-2 x AAT95785 ..

Align seg 1/1 to: AAT95785 from: 1 to: 2383

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20

41 TCGCCGCGTGGCGCGCGTCCCTC..... 67

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34

```

  ||| ||||| ..... |||||
68 .CTCGCGCGCGCGCTGCGCGCGCGCTGCGCGACCGCGCGCGCGCCGC 116
35 .ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
   |||||:||||: ||| ||| |||||:|||||
117 GCGCGCTCCAGCGCGTCTGCGCGCGCGTGCCTGCGGAGCTGAGCAGG 166
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
   |||||:||||| |||||
167 GAGGCGCGCGCGCGCGTCCACCGCGCTGCTGGCGCGCGCGCT 216
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
   |||||:||||| ||||| |||||:|||||
217 CGTCCCGCGCTCTCTCGCG...CCGCGCGCGAGCCCGCGGTGAGCGG 263
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
   ||| |||||:||||| |||||:||||| |||||
264 CATTGACGCGCGCGCGTCCCGCGCGCGCTGGGGGTCTCGGTGTC 313
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
   :|| |||||:||||| |||||:||||| :|||
314 GAA.....CCTGAAGGATTGCTGAAGGTTCATCGATAACACAGTAGT 357
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
   |||||:||||| |||||:||||| :|||
358 TGTGGCAAGTGAGCAAGATTCTGAGATTGTTGGTGGAAAGGAGCAAGCTC 407
134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
   :||| |||||:||||| |||||:||||| :|||
408 GAGCTAAAGTAACACAAACATTTCTTTGTAAGTGGCGAAGCTTCTCCT 457
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
   |||||:||||| |||||:||||| |||||:|||||
458 TATGCAAAAGCTGGGGGTCTAGGAGATGTTTGTGTTTCATTGCCAGTTGC 507
167 aLeuAlaAlaArgGlyHisArgValMetValMetProArgTyrLeuA 184
   |||||:||||| |||||:||||| |||||:|||||
508 TCTTGCTGCTGGTCCACCGTGTGATGTTGTAATCCCGAGATATTTAA 557
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
   |||||:||||| |||||:||||| |||||:|||||
558 ATGTGTACCTCCGATAAGAAATTATGCAAAATGCAATTTTACAGAAAAACAC 607
201 lLelysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
   |||||:||||| |||||:||||| |||||:|||||
608 ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATGA 657
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
   |||||:||||| |||||:||||| |||||:|||||
658 GTATAGAGATTCAGTTGACCTGGGTGTTTGTGATCATCCCTCATATCACA 707
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
   |||||:||||| |||||:||||| |||||:|||||
708 GACCTGAAAAATTTATATGGAGATAAGTTGGTGTCTTTGGTGAATAATCAG 757
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
   |||||:||||| |||||:||||| |||||:|||||
758 TTCAGATACACACTCCTTGTCTGCTGCAATGAGGCTCCTTTGGTCTCT 807
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
   |||||:||||| |||||:||||| |||||:|||||
808 TGAATTGGGAGGATATATTTATGGACACAAATTCATGTTTGTGTCATG 857
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
   |||||:||||| |||||:||||| |||||:|||||
858 ATTCGCATGCCAGTCTAGTCCAGTCTCTTGTGTCGCAAAATATAGACCA 907
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
   |||||:||||| |||||:||||| |||||:|||||
908 TATGCTGTTTATAAGACTCCCGCAGCATCTTGTGTAATACATAATTTAGC 957
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
   |||||:||||| |||||:||||| |||||:|||||

```

```

958 ACATCAGGGTGTAGAGCTGCAAGCACATATCCTGACCTTGGGTGGCCAC 1007
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaIleArgArg 350
|||||
1008 CTGAATGGTATGGAGCTCTGGAGTGGGTATTCCTCGAATGGGGAGGAGG 1057
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||
1058 CATGCCCTTGACAAGGGTGAGCGAGTTAAATTTTGAAGAGTGCAGTTGT 1107
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
|||||
1108 GACAGCAGATCGAATCGTGAATGTCAGTAAGGGTTATTTCATGGGAGGTCA 1157
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||
1158 CAACCTGCTGAAGGTGGACAGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1207
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||
1208 AGTGATTTAAACCGAATTTGTAATGGAATTTGACATTAATGATGGAACCC 1257
417 oThrThrAspLysCysLeuProHisLysTyrSerValAspAspLeuSerG 434
|||||
1258 TGCCACAGACAATGATCCCTGCTCATTTATCTGTGATGACCTCTCTG 1307
434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuLeuProVal 450
|||||
1308 GAAAGGCCAAATGTAAGGTGCATTCAGAAAGGAGCTGGGTTTACCTATA 1357
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
|||||
1358 AGCCCTGATGTTCCCTCTGATGGCTTTATTTGGAAGATTGGATTATCAGAA 1407
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||
1408 AGGCATTGATCTCATTCACATTCATACACAGATCTCATCGCGGAAGATG 1457
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||
1458 TTCAATTTGTCATGCTGTGATGCTGTGTGACCCAGAGCTTGAAGATTGGATG 1507
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
|||||
1508 AGATCTACAGATCGATCTTCAAGGATAAATTCGTGGATGGGTGGATT 1557
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
|||||
1558 TAGTGTTCAGTTTCCACCGAATAAATCGCCGCTCGCATATATTGTTAA 1607
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||
1608 TGCCATCCAGATTCGAACCTTGTGCTCTCAATCAGCTATATGCTATGCAG 1657
551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||
1658 TATGCGACAGTTCCCTGTTGTCATCCACTGGGGCCCTTAGAGATACCCT 1707
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584
|||||
1708 GGAGAACTTCAACCTTTCGGTGAAGATGGAGAGCAGGTTACAGGTTGGG 1757
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||
1758 CATTGCGACCCCTAACCCAGAAAACATGTTGTGGACATTCGGAACCTGCA 1807
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||
1808 ATATCTACATACAGGGAACACAAAGTCTCTGGGAAGGCTTAATGAAGCG 1857
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSer 633
|||||
1858 AGGCATGTCAAAAGACTTCACGTGGGACCATCGCGCTGAACAAATACGAAC 1907

```

```

634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
|||||
1908 AAATCTTCCAGTGGGCTTCATCGATCGACCTTATGTCATGTAATAAAGG 1957
650 yLeuGlyArgSerLysCys.GluSerProSerSerAlaLeuLysThrSerSer 666
|||||
1958 A.....CCAAAGTGGTTCCTTGAAGATCATCAGTTTCATCATCC 1998
667 SerSerPheArgGlyProGluGlyTyrPro.....CysThrLeuArgCy 681
|||||
1999 TATAGTAAGCTGAATGATGAAGAAACCCCTGATACATTACATGGAAGC 2048
681 sProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlys 698
|||||
2049 AGACCGGCTATTGGCTCCATTGCTCCAATGCTGCTTTGGCTGCCT... 2094
698 erArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGly 714
|||||
2095 .....TGCTCGATGGACCGGATGCGATGAGGAATCCA 2127
715 ArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla.....Al 728
|||||
2128 GCCGAACGACAGTCTTGAAGGATAGGAAGGGAGCTGGAAGCAGTCACGC 2177
728 aGlyTrpLeuThrAlaLysHisSerAspGlySerLeuSer 742
|||||
2178 AGGCAGCTCGCCGTTTCATATATGGAACAGCTGGAGTCACT 2220

```

seq_name: /SIB1/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50643

seq_documentation_block:

ID_AAZ50643 standard; cDNA; 2008 BP.

AC_AAZ50643;

23-MAY-2000 (first entry)

Corn soluble starch synthase gene fragment inserted in pSS65-C11.

Soluble starch synthase; starch fine structure; corn;
transgenic plant; amylose; amylopectin; amylose polymerisation;
non-granule bound starch synthase; non-GBSSI; altered starch; food;
paper; plastic; adhesive; ss.

Zea mays.

WO200006755-A2.

10-FEB-2000.

26-JUL-1999; 99WO-US16296.

28-JUL-1998; 98US-0094436.

(DUPO) DU PONT DE NEMOURS & CO E I.

Broglie KE, Lightner JE;

WPI; 2000-195311/17.

Producing transgenic cereal crops with altered starch structure useful
for preparing foodstuff, paper, plastic or adhesives, comprises
transforming crops with chimeric sense or antisense gene construct
encoding starch synthase

Claim 5; Page 53; 56pp; English.

The present sequence is the corn soluble starch synthase (SSI) DNA
sequence comprising the entire SSI coding region and a 3' UTR fragment
inserted into plasmid pSS65-C11. The chimeric gene containing the zein
promoter followed by the 3'UTR is used as a sense construct for
preparation of transgenic corn expressing altered starch structure. The

228 sPHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPhGly 244
 |||||
 545 ATCATCCTCATATACAGACCTGGAAATTTATATGGAGATAAGTTGGT 594
 245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystTyrAlaAlaCys 261
 |||||
 595 GCTTTTGGTGATAATCATAGTTTCAGATACACACTCCTTTGCTATGCTGCATG 644
 261 sGluAlaProLeuLeuLeuLeuGlyGlyTyrIleTyrGlyGlnAsnC 278
 |||||
 645 TGAGGCTCCTTTGATCCTTGAATGGGAGGATATATTTATGGACAGAATT 694
 278 yMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294
 |||||
 695 GCATGTTTGTTCATAGTATGGCATGCCAGTCTAGTGCAGCTTCCT 744
 295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311
 |||||
 745 GCTGCAAAATATAGCATATGTTGTTTATAAAGACTCCCGCAGCATTC 794
 311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328
 |||||
 795 TGTAAATACATAATTTAGACATAGGGGTAGAGCCTGCAGACCATATC 844
 328 roAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGluTyrValPhe 344
 |||||
 845 CTGACCTTGGGTTGCCACCTGAAATGGTATGGAGCTCTGGAGTGGGTATTC 894
 345 ProGluTyrAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361
 |||||
 895 CCTGAATGGCGAGGAGGAGCATGCCCTTTGACAAGGGGTGAGGCGTAAATTT 944
 361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnG 378
 |||||
 945 TTTGAAAGGTGCAGTTGACAGCAGATCGAATCTGACTGCTCAGTAAGG 994
 378 lyTyrSerTrpGluValThrThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394
 |||||
 995 GTTATTCGTGGAGGTCAACAAGTCTGAAGGTGGACAGGCGCTCAATGAG 1044
 395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411
 |||||
 1045 CTCTTAAGCTCCAGAAAGAGTGTATTAACCGGAATGTAAATGGAATGA 1094
 411 pileAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428
 |||||
 1095 CATTAATGATGGAAACCTGCCACAGACAATGTATCCCTGTCATTATT 1144
 428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
 |||||
 1145 CTGTTGATGACCTCTCTGGAAGGCCAAATGTAAAGGTGCATTCGAGAG 1194
 445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG 461
 |||||
 1195 GAGCTGGGTTTACCTATTAAGGCTGATGTTCTCTGATTTGGCTTTATGG 1244
 461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478
 |||||
 1245 AAGTTGGATATACAGAAAGGCAATGATCTCACTCAACTTATCATACCAG 1294
 478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494
 |||||
 1295 ATCTCATCGCGGAGAGATGTTCAATTTGTCATGCTTGGATCTGGTGACCCA 1344
 495 IlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspLysPh 511
 |||||
 1345 GAGCTTGAAGATGGATGAGATCTACAGATCGATCTTCAAGGATTAAT 1394
 511 eArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAlaG 528
 |||||
 1395 TCGTGGATGGGTTGATTTAGTGTCCAGTTTCCACCAGCAATAACTGCCG 1444
 528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544

1445 GCTGCGATATATTGTTAATGCCATCCAGATTCGAACCTTGTGGTCTCAAT 1494
 545 GlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrG 561
 |||||
 1495 CAGCTATATGCTATGCAGTATGCACAGTTCCTGTTGCTCCATGCCACTGG 1544
 561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578
 |||||
 1545 GGGCCTTAGAGATACCGTGGAGAACTTCAACCTTTCCGTTGAGAAATGGAG 1594
 578 luGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLe 594
 |||||
 1595 AGCAGGGTACAGGGTGGCATTCGCACCCCTAACCCCTTTCCGTTGAGAAATGGAG 1644
 594 uTTPAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSer 611
 |||||
 1645 CTGGACATTCGGAACGTGCAATATCTACATACAGGGAACACAACTCCTCCT 1694
 611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627
 |||||
 1695 GCGAAGGGCTAATGAAGCGAGGATGTCAAAGACTTCACGTGGGACCAT 1744
 628 Ala 628
 |||||
 1745 GCC 1747
 seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV70960
 seq_documentation_block:
 ID AAV70960 standard; DNA; 1749 BP.
 XX
 AC AAV70960;
 XX
 DT 23-AUG-1999 (first entry)
 XX
 DE DNA encoding maize starch soluble synthase 1-2..
 XX
 KW Non-glycogen-like polysaccharide production; fermentation;
 KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;
 KW non-starch branching gene; amylopectin; amylose; plant-like starch;
 KW maize starch soluble synthase 1-2; ss.
 XX
 OS Zea mays.
 XX
 PN WO9844780-A1.
 XX
 PD 15-OCT-1998.
 XX
 PF 03-APR-1998; 98WO-US06660.
 XX
 PR 04-APR-1997; 97US-0042939.
 XX
 PA (EXSE-) EXSEED GENETICS LLC.
 XX
 PI Guan H, Keeling PL;
 XX
 DR WPI; 1998-568285/48.
 DR P-PSDB; AAV70894.
 XX
 PT Producing non-glycogen-like polysaccharides in bacteria, fungi or
 PT plants - transformed with genes for enzymes involved in starch or
 PT glycogen synthesis allows fermentative production of starches with
 PT engineered properties
 PS Disclosure; Fig 52; 150pp; English.
 CC
 CC The specification describes a method for the production of
 CC non-glycogen-like polysaccharides in a host. The method comprises
 CC transforming a host, suitable for fermentation, with genes encoding
 CC starch- or glycogen-synthesis enzymes, and fermenting the transformants.
 CC The specification also describes hosts transformed with a gene active
 CC in glycogen synthesis and at least one non-starch branching gene,

Involved in production of amylopectin or amylose in its original host. The method is used to produce plant-like starches by fermentation and new starches in plants. These starches are useful for all food and non-food applications of starch. The present sequence is used in the course of the invention.

SQ Sequence 1749 BP; 432 A; 378 C; 469 G; 470 T; 0 other;

alignment_scores:	
Quality:	2708.50
Ratio:	4.916
Percent Similarity:	94.188
Percent Identity:	86.154
Gaps:	3
Length:	585

alignment_block: .
US-09-674-824-2 x AAV70960

Align seg 1/1 to: AAV70960 from: 1 to: 1749

45 ValAlaGluLeuSerArgGluGlyProAlaIleAraGrProAlaGlnGlnG1 61
4 GTCCGGAGCTGAGCAGGAGGGCGCGGCCGCCTGCCACCCGC 53
61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78
54 GCTGTGGCGCCCGCTGTGCCGGGTTCCTCGC...CGCGCGCGCG 100
78 laProAlaGlnSerProAlaProThrGlnProProLeuProAspAlaGly 94
101 AGCCCAAGGTTGAGCGGCATCGACGCCCGCGTGGCCGACGCCGCG 150
95 ValGlyGluLeuAlaProAspLeuLeuGluGlyIleAlaGluAspSe 111
151 CTGGGGACCCTCGGT....CTCGAACCTGAAGGATTGCTGAAGTTC 194
111 rIleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMetA 128
195 CATCGATAACACAGTAGTGTTGGCAAGTGACCAAGATTCTGAGATGTGG 244
128 spAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPheVal 144
245 TTGGAAAGGACCAAGCTCGACTTAAGTTAACACAAGAAGCATGTCTTTGTA 294
145 ThrGlyGluAlaAlaProTyrrAlaLysSerGlyGlyLeuGlyAspValcy 161
295 ACCGCGAAGCTTCTCCTTATGCAAAGTCTGGGGTCTAGGAGATGTTG 344
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisrgValMetValV 178
345 TGGTTATTCGCAAGTGTCTGTGTGCTGCTGGTCCACCGCTGATGTTG 394
178 alMetProArgTyrrLeuAsnGlySerSerAspLysAsnTyrrAlaLysAla 194
395 TAATGCCACAGATATTAATGGTACCTCCGATAGAAATATTGCAATGCA 444
195 LeuTyrrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisGl 211
445 TTTTACACAAAAACACATTCGGATTCCATGCTTTGGCGGTGAACATGA 494
211 uValThrPhePheHisGluTyrrArgAspAsnValAspTrpValPheValA 228
495 AGTTACCTCTTCATCAGTATAGAGATTACAGTTGACTGGGTGTTGTTG 544
228 sphisProSerTyrrHisArgProGlySerLeuTyrrGlyAspAsnPheGly 244
545 ATCATCCTCATATCACAGACCTGGAATTTATATGAGATGAAGTTGGT 594
245 AlaPheGlyAspAsnGlnPheArgTyrrThrLeuLeuCystyrAlaAlaCy 261
595 GCTTTTGGTGATAATCAGTTTCAGATACACACTCCTTTGCTATGCTGCATG 644
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrrIleTyrrGlyGlnAsnC 278

645 TGAGCGTCCTTTGATCCTTGTAATTTGGGAGGATATATTTATGCACAGAATT 694

278 ysMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294

695 GCATGTTTGTGTCAATGATTGGCATGCCAGCTCTAGTGCACAGTCTTCTT 744

295 AlaAlaIysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311

745 GCTGCAAAATATAGACCATATGGTGTATTATAAGAGACTCCCGCAGCATCT 794

311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328

795 TGTAAATACATAATTTACACACATCAGGGTGTAGACCTGCAAGCACATATC 844

328 roAspLeuGlyLeuProProGluTrpTyrGlyAlaLeuGluTrpValPhe 344

845 CTGACCTTGGGTTGCCACCTGAATGGTATGGAGCTCTGGAGTGGGTATTC 894

345 ProGluTrpAlaArgArgHisAlaLeuAspIysGlyGluAlaValAsnPh 361

895 CCTGAATGGCGAGGAGCATGCCCTTGACAAGGTCAGGCGATTAATTT 944

361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGln 378

945 TTTGAAAGGTCAGTGTGTGACAGCAGATCGAATCGTGACTGTCAGTAAGG 994

378 lyTyrSerTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394

995 GTTATTGCTGGGAGGTCACAACTGCTGAAGGTGCACAGGCGCTCAATGAG 1044

395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAlaAsnGlyIleAs 411

1045 CTCTTAAGCTCCAGAAAGAGCTGATTAAACGGAAATGTAAATGGAATTGA 1094

411 pIleAsnAspTrpAsnProThrThrAspLysCysLysLeuProHisHisTyrS 428

1095 CATTAATGATTGGAACCTGCCACAGACAAATGATGCCCTGTCATTAATT 1144

428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444

1145 CTGTTGATGACCTCTCTGGAAAGGCCAAATGTAAGGTCGATTCGACGAAG 1194

445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleGl 461

1195 GAGCTCGGGTTTACCTATAAGCCCTGATGTTCTCTGATTGGCTTTATTGG 1244

461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478

1245 AAGGTTGGATTATCAGAAAGCATTCATCTCATTTCACTTATCATACCAG 1294

478 luLeuMetArgGluAspValGlnPheValImetLeuGlySerGlyAspPro 494

1295 ATCTCATCGGGAAGATGTTCAATTTGTATGCTTGGATCTGGTGACCCA 1344

495 IlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspLysPh 511

1345 GAGCTTGAAGATTGGATGAGATCTCAGAGTCGATCTTCAAGGATTAATTT 1394

511 eArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAlaG 528

1395 TCGTGGATGGGTTGGATTTAGTGTTCACGTTTCCCAACCGAATAACTGCCG 1444

528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544

1445 GCTGCGATATATGTTAATGGCATCCAGATTCGAACCTTGTGGTGCATCAAT 1494

545 GlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrGl 561

1495 CAGCTATATGCTATGCAGTATGSCACAGTTCCTGTTGTCATGCACACTGG 1544

561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578

1545 GGGCCCTTAGAGATACCGTGGAGAACTCAACCCCTTTCGGTGAAGATGGAG 1594

578 luGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMet.Le 594
|||||
1595 AGCAGGTACAGGTGGGCATTCACCCCTTACCACAGAAACATGTTT 1644
594 utrAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611
|||||
1645 GTGCACATTGGGAACCTGCAATATCTACATACAGGGAACAAAGTCTCTCT 1694
611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627
|||||
1695 GGGAGGGCTAATGAAGCGAGGCAATGTCAAAGACTTTCACGTGGGACCAT 1744
628 Ala 628
1745 GCC 1747

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.AAT67285

seq_documentation_block:

ID AAT67285 standard; cDNA; 2992 BP.

XX AAT67285;

XX 11-SEP-1997 (first entry)

DE Soluble starch synthase cDNA clone SSS10.52.

XX Soluble starch synthase; SSS10.52; transgenic plant; cereal; maize;
KW ss.

OS Zea mays inbred line W64A.

PN W09720936-AL.

XX 12-JUN-1997.

XX 04-DEC-1996; 96WO-GB02990.

XX 06-DEC-1995; 95GB-0024938.

XX (ZENE) ZENECA LTD.

XX Keeling PL, Knight ME;

XX WPI; 1997-319782/29.

XX cDNA encoding soluble starch synthase - used to produce transgenic
PT plants with increased capacity for producing and storing starch

XX Claim 5; Page 17-20; 44pp; English.

XX cDNA clones SSS10.52, SSS6.31 and SSS56 (AAT67285-87) code for maize
CC soluble starch synthase (SSS). They were isolated from a maize
CC inbred line W64A library by screening with a probe (AAT67288) based
CC on pea SSS, rice SSS and maize GBSS. The cDNAs can be used to
CC produce plants with an increased capacity for producing starch, or
CC a capacity to produce starch with an altered fine structure. They
CC can also be used to isolate the corresponding genomic sequences
CC from crop plants, to determine the contribution of the SSS gene to
CC the net regulation of starch biosynthesis, and to modify the levels
CC of starch produced by the plant. Transgenic plants (esp. maize)
CC can be used to produce hybrid plants which have higher rates of
CC starch synthesis at temperatures above the normal optimum.

XX Sequence 2992 BP; 758 A; 655 C; 801 G; 776 T; 2 other;

alignment_scores:

Quality: 2655.50 Length: 745
Ratio: 4.304 Gaps: 11
Percent Similarity: 82.819 Percent Identity: 71.544

alignment_block:

US-09-674-824-2 x AAT67285 ..

Align seg 1/1 to: AAT67285 from: 1 to: 2992

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
614 TCGGCCGTGGCGCCGCGTGCCTC.....CTCTCCGCGG 648
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAlaArgLeuA 37
649 GNGCGCTGCCCGCCGCGCTCGCGACCGCGCGCGCGCGAGGCTCC 698
37 rgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPro 53
699 AGCGCGTGCCTCGCGCGCGCGTGCCTCGCGAGCTGAGCAGGGAGGGGCC 748
54 AlaAlaArgProAlaGlnGlnGlnAlaProProLeuValProGl 70
749 GCGCGCGCGCGCTGCCACCGCGCTGCTGCCCGCGCGCTCGTCCCGG 798
70 yPheLeuAlaProProProAlaProAlaGlnSerProAlaProThrG 87
799 CTTCCTCGCG...CCGCGCGCGAGCCACGGGTGAGCCGCGCATCGACGC 845
87 lnProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeu 103
846 CGCGCGCGCGCGCGCGCTGGGGACCTCGGT.....CTCGAA 889
104 LeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaAla 120
890 CCTGAAGGGATTGCTGAAGGTTCCATCGATAACACACTAGTTGTGGCA 939
120 rGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAlaLys 137
940 TGACCAAGATTCTCAGATTGTGTTGGAAGGAGCAAGCTCGAGCTAAG 989
137 alThrArgSerIleValPheValThrGlyGluAlaAlaPro.TyrAlaLy 153
990 TAACACAAGCATTGTCTTTGTAACCGCGGAAGCTTCTCTTAATCGAAA 1039
153 sSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAla 170
1040 GTCTGGGGTCTAGGAGATTGTTGTTGTTTCCAGTTGCTCTTGTCTG 1089
170 laArgGlyHisArgValMetValValMetProArgTyrLeuAsnGlySer 186
1090 CTCGGGTTCACCGTGTGATGTTGTAATGCCCGACACATTTAAATGGTACC 1139
187 SerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLys 203
1140 TCCGATAAGAATTATGCAATGCATTTTACTCAGAAAAACACATTCGGAT 1189
203 eProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgA 220
1190 TCCATTCTTGGCGGTGAACATGAAGTTACCTTCTTCCATGAGTATAGAG 1239
220 spAsnValAspTrpValPheValAspHisProSerTyrHisArgProGly 236
1240 ATTCAGTTGACTGGGTGTTTGTGATCATCCTCATATCACAGACCTGGA 1289
237 SerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTy 253
1290 AATTTATATGAGATAAGTTTGGTCTTTTGGTGATAATCATGTTTCAGATA 1339
253 rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuG 270
1340 CACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGCTTGTGAATTGG 1389
270 lyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHis 286
1390 GAGGATATATTTATGACACAGAATTGCTGTTGGTTGTCATGATGGCAT 1439

```
287 AlaSerLeuValProValLeuAlaAlaLysTyrArgProTyrGlyVa 303
1440 GCCAGTCTAGAGCCAGTCTCTTGTGCTGCAAAATATAGCCATATGGTGT 1489
303 lTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnG 320
1490 TTATAAGACATCCCGCAGCATCTTGTAAATACATAAATTAGCACATCAGG 1539
320 lYValGluProAlaSerThrTyrProAspLeuGlyLeuProProGluTrp 336
1540 GTGTAGACCTCGCAAGCACATATCTGACCTTGGGTTGGCCACCTGAATGG 1589
337 TyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLe 353
1590 TATGGAGCTCTGGAGTGGGTATTCCTGCAATGGCGAGGAGCATGGCCT 1639
353 uAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrAlaA 370
1640 TGACAAGGGTGAGGCAGCTTAATTTTTTGAAGAGTGCAGTTGTGACAGCAG 1689
370 sPArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThrAla 386
1690 ATCGAATCGTCACTGTCAAGTGGTATTCATGGGAGTCCAACTGCCT 1739
387 GluGlyGlnGlyLeuAsnGluLeuSerSerArgTysSerValLe 403
1740 GAAGGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGATT 1789
403 uAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThra 420
1790 AAACGGAAATGTAATGAATGACATTAATGATTGGAACCCCTGCCACAG 1839
420 sPLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAla 436
1840 ACAAAATGATATCCCTGTCTATTCTGTGTGATGACCTCTCTGAAAGGCT 1889
437 LysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAs 453
1890 AAATGTAAGGTGTCATTTCGCAAGAGGAGCTGGGTTTACCTATAAGGCCTGA 1939
453 pValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleA 470
1940 TGTTCCTCTGATGGCTTTATTGGAAGATTGGATTATCAGAAAGCAATTG 1989
470 sPLeuIleLysMetalIleProGluLeuMetArgGluAspValGlnPhe 486
1990 ATCTCAATCAACTATCATACCATCATCGGCGAAGAAATGTTCAATTT 2039
487 ValMetLeuGlySerGlyAspProIlePheGluClyTrpMetArgSerTh 503
2040 GTCATGCTTGATCTGTGACCCAGAGCTTGAAGATTGGATGAGATCTAC 2089
503 rGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValp 520
2090 AGATTCGATCTTCAAGGATAAATTCGTGGATGGGTTGGATTAGTGTTC 2139
520 roValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSer 536
2140 CAGTTTCCCAACCGAATAACTCGCGCTGGCGATATATTGTTAATGCCATCC 2189
537 ArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyTh 553
2190 AGATTGCAACCTTGTGTCTCAATCAGCTATATGCTATGCAAGTATGGCAC 2239
553 rValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrP 570
2240 AGTTCCTGTTGTCATGCACTGGGCGCTTACAGATACCGCTGGAGAACT 2289
570 heAsnProPheGlyAlaLysGlyGluGlyThrGlyThrTrpAlaPheSer 586
2290 TCAACCCCTTTCGGTGAAGATGGAGAGCAGGGTACAGGGTGGGCATTCGCA 2339
587 ProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerT 603
```

```
2340 CCCCTAACCCAGAGAAACATGTTTGTGACATTCGAACTGCAATATCTA 2389
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
2390 CATACAGGGAACACAAAGTAATAATGGGAAGGCTAATGAAGCCAGGCATG 2439
620 ThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSers 636
2440 TCAAAAGAGTTTCAGTGGGACCATGCCCTGGAACAATACGAACAATCTT 2489
636 erSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG1 652
2490 CCATGGCCCTTCATCGGATCGACCC...GATGTTCAATGAAAAAGGG 2536
652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerp 669
2537 ACCAAAGTGTGTTGTTCCCTTGAAGATTATCAGTTTCATCATCTTATAGTA 2586
669 heArgGlyProGluGlyTyrPro...CysThrLeuArgCysProAla 683
2587 AGCTGAATGATGAAGAAACCCCTGTACATTACATGGAAGGCAGACCGG 2636
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPhe 695
2637 CTATTGGCTCCATTGCTCCCAATGCTGCTTGGCTTGGCTTCCCTCGCATGG 2686
695 2687 ACCGGATGTCAGTGAGGAATCCAGNCGAACGACAGTTTTGAAGGATAGGAA 2736
696 .....AlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThr 710
2737 GGGAGCTGGAAGC...AGTCACGCGAGGCGAGGCAAGCCTTCGCCGTTAAT 2783
711 AlaSerGlyGlyArgGlnLeuGlnPhe 719
2784 TCATATGGAACAAGCTGGAGTCACTTT 2810
```

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.AAV66832

seq_documentation_block:

ID AAV66832 standard; DNA; 2990 BP.

XX AAV66832;

XX AC

XX DT 05-JAN-1999 (first entry)

XX DE Zea mays soluble starch synthase gene SSS1052 and SSS64.

XX KW Zea mays; US yellow-dent corn line; maize; soluble starch synthase;

XX KW SSS; glycogen biosynthetic pathway; branching enzyme; ss.

XX OS Zea mays.

XX PN US5824790-A.

XX PD 20-OCT-1998.

XX PF 15-DEC-1995; 950S-0572951.

XX PR 15-DEC-1995; 950S-0572951.

XX PR 21-JUN-1994; 94US-0263921.

XX PR 29-NOV-1994; 94US-0346602.

XX XX (ZENE) ZENECA LTD.

XX PI Guan H, Keeling PL, Knight ME;

XX DR WPI; 1998-582626/49.

XX PT Isolated nucleic acid molecule, used to produce transgenic plants - comprises nucleotide sequence encoding polypeptide having soluble


```

||||| .....
2092 GTCGATCTTCAAGGATAAAATTCGTGGATGGTGTGATTTAGTGTCCAG 2141
521 alSerHisArgIleThrAlaGlyCysAspIleLeuMetProSerArg 537
||||| .....
2142 TTTCCACCAAGTAATACTGGGCTGGGATATATTCTTAATGCCATCCAGA 2191
538 PheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrVa 554
2192 TTCGAACCTTGGTCTCAATCAGCTATATGCTATGCAGTATGGACAGT 2241
554 lProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheA 571
||||| .....
2242 TCCTGTGTGCCATGCAACTGGGGCTTAGAGATACCGTGGAGAACTTCA 2291
571 snProPheGlyAlaLysGlyGluGluGlyThrGlyTyrAlaPheSerPro 587
||||| .....
2292 ACCCTTCCGTGAGATGGAGACGAGGTACAGGTGGCATTCGCACCC 2341
588 LeuThrValAspLysMet.LeuTyrAlaLeuArgThrAlaMetSerThrP 604
||||| .....
2342 CTAACACAGAAACATGTTGTGGACATTCGAACTGCAATATCTACAT 2391
604 heArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThr 620
||||| .....
2392 ACAGGGAACACAGTAATAATGGGAAGGGCTAATGAAGCCAGGATGTCA 2441
621 LysAspHisThrTrpAspHisAlaPro.SerSerThrSerArgSerSerS 637
||||| .....
2442 AAAGATTCACGTGGGACCATGCCCTCAACAATACGACAAATCTTCCA 2491
637 erGlyProSerTrpThrAsnPro 644
||||| .....
2492 GTGGCCCTTCATCGATCGACCC 2514

```

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.AAV29759

seq_documentation_block:

ID AAV29759 standard; DNA; 1620 BP.

AC AAV29759;

11-SEP-1998 (first entry)

Zea mays pEXS52 starch synthase gene.

SER; starch-encapsulating region; fusion vector;
 starch synthase; glucosyl transferase; pEXS52; ss.

Zea mays.

Key	Location/Qualifiers
FT CDS	1..2007
FT	/*tag= a
FT	/product= starch synthase

W09814601-Al.

09-APR-1998.

30-SEP-1997; 97WO-US17555.

30-SEP-1996; 96US-0026855.

(EXSE-) EXSEED GENETICS LLC.

Guan H, Keeling P;

WPI; 1998-240100/21.

P-PSDB; AAW56491.

Hybrid polypeptide comprising starch-encapsulating region and
 protein - useful for, e.g. producing protein(s) resistant to

PT degradation by stomach acids

XX Example 5; Page 51-53; 156pp; English.

The sequence is that of the starch synthase coding region from
 pEXS52. It can be used in the production of a hybrid polypeptide
 comprising a starch-encapsulating region (SER) fused
 to a payload protein. The hybrid polypeptide can be used to make
 modified starches comprising the payload protein, selected from,
 e.g. hormones, growth factors, antibodies, enzymes, dyes,
 immunoglobulins, etc. The modified starch can also be used
 to provide grain feeds enriched in amino acids. By encapsulating
 the payload protein in starch, it is more resistant to
 degradation by stomach acids.

XX Sequence 1620 BP; 426 A; 313 C; 422 G; 459 T; 0 other;

alignment_scores:

Quality:	2546.50	Length:	538
Ratio:	4.935 <td>Gaps:</td> <td>3</td>	Gaps:	3
Percent Similarity:	95.911 <td>Percent Identity:</td> <td>87.918</td>	Percent Identity:	87.918

alignment_block:

US-09-674-824-2 x AAV29759 ..

Align seg 1/1 to: AAV29759 from: 1 to: 1620

```

95 ValGlyGluLeuAla...ProAspLeuLeuLeu.....GluGlyIleAl 108
||||| .....
4 GTCGGGAGCTGAGCAGGAGGACCTCGGTCTCGAACCTGAAGGGATTGC 53
108 aGluAspSerIleAspSerIleValAlaAlaSerGluGlnAspSerG 125
||||| .....
54 TGAAGGTTCCATGCAACACAGTAGTTGTGCAAGTGAGCAAGATTCTG 103
125 lIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSerIle 141
||||| .....
104 AGATTGGTGTGGAAGGAGCAGCTCGAGCTAAAGTAACACAAAGCAT 153
142 ValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuG 158
||||| .....
154 GTCTTTGTAACCGCGGAAGCTTCTCTTATGCAAAAGCTGGGGGTCTAG 203
158 yAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArg 175
||||| .....
204 AGATTTGTGGTTCATTGCCAGTTGCTCTGCTGCTGCTGCTGCTGCTG 253
175 aMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyr 191
||||| .....
254 TGATGTTGTAATGCCAGATATTTAAATGGTACCTCCGATGAAGATTAT 303
192 AlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyG 208
||||| .....
304 GCAAAATGCATTTTACACAGAAAAACACATTCGATTCCATGCTTTGG 353
208 ySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrp 225
||||| .....
354 TGAACATGAAGTTACCTTCTCCATGAGTATAGAGATTGAGTTGAGTGG 403
225 aPheValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAsp 241
||||| .....
404 TGTTTGTTGATCATCCCTCATATACAGACCTGGAAATTTATATGAGAT 453
242 AsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCys 258
||||| .....
454 AAGTTTGGTGTGTTTGGTGATAATCAGTTTCAGATACACACTCCTTTG 503
258 rAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyr 275
||||| .....
504 TGCTGCATGTGAGGCTCCTTTGATCTTGAATGGGAGGATATATTTATG 553
275 lGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuValPro 291

```

Fri Mar 29 09:26:43 2002

```
|||||
554 GACAGAATTGCATGTTTGTCTCAATGATTGGCATGCCAGCTAGTGCCA 603
292 ValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerAr 308
|||||
604 GTCCTCTCTGCTGCAAAATATAGACCATATGTTGTTTATAAGACATCCCG 653
308 gSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValIleProAlas 325
|||||
654 CAGCATCTCTGTAATACATAAATTTAGCACATCAGGGGTGTAGAGCCTCAA 703
325 erThrTyrProAspLeuGlyLeuProGluTyrProGluTyrGlyAlaLeuGlu 341
704 GCACATATCTCGACCTTGGGTGGCAGCTGAATGATGGAGCTCTGGAG 753
342 TrpValPheProGluTyrAlaArgHisAlaLeuAspLysGlyGluAl 358
754 TGGGTATTCCTGAATGGCGAGGAGCATGCCCTTGACAAGGGTGAGGC 803
358 aValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrV 375
804 AGTTAAATTTTGAAGGTCAGTGTGTGACAGCAGATCGAATCGTGACTG 853
375 alSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGlnGly 391
854 TCAGTAGGGTTATTCGTGGAGGTCACACTGCTGAAGGTGGACAGGC 903
392 LeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAs 408
904 CTCATAGAGCTCTTAAGCTCCAGAAGAGTGTATTAAACGGAATGTAA 953
408 nGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuProH 425
954 TGAATTTGACATTAATGATTGAACCTTGGCCACAGACAAATGTATCCCT 1003
425 lSHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGlu 441
1004 GTCATTATTCTTGTGATGACCTCTCTGGAAGGCCAAATGTAAGGTGCA 1053
442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleGl 458
1054 TTGCAGAGAGGAGCTGGGTTTACATATAGGCTGATGTTCTCTCTGATGG 1103
458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMeta 475
1104 CTTATTGGAAGTTGGATTATCAGAAAGGCATTGATCTCATTTCAACTTA 1153
475 laileProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491
1154 TCATACACAGATCTCATGCGGGAAGATGTTCAATTTGTCATGCTTGGATCT 1203
492 GlyAspProIlePheGluGlyTrpMetArgSerThrGluSerTyrIly 508
1204 GGTGACCCAGAGCTTGAAGATTGGATGGATGATGATGATGATGATGATG 1253
508 sAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArgI 525
1254 GGATAAATTTCTGGATGCTGGATTTAGTGTTCAGGTTTCCACCCGAA 1303
525 leThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCys 541
1304 TAACTCCGCGCTCGCATATATTGTTAATGCCATCCAGATTCGAACCTTGT 1353
542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValH 558
1354 GGTCTCAATCAGCTATATGCTATGATGATGATGATGATGATGATGATG 1403
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575
1404 TGCAACTGGGGGCTTAGAGATACCGTGGAGAACTTCAACCCCTTTGCGTG 1453
575 laLysGlyGluGluGlyThrGlyTyrAlaPheSerProLeuThrValAsp 591
|||||
```

```
1454 AGAATGGAGCAGCGGTACAGGGTGGGCATTCGCACCCCTCAACACAGAA 1503
592 LysMet. LeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisL 608
|||||
1504 AACATGTTTGTGACATTCGGAACATGCAATATCTACATACAGGAAACACA 1553
608 ysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 624
1554 AGTCTCTCTGGGAAAGGCTAATGAACGAGGATGTCAAAGACTTTCAGC 1603
625 TrpAspHisAla 628
1604 TGGGACCATGCC 1615
seq_name: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50637
seq_documentation_block:
ID AAZ50637 standard; cDNA; 1528 BP.
XX
AC AAZ50637;
XX
DT 23-MAY-2000 (first entry)
XX
DE Corn soluble starch synthase gene fragment inserted in pSS42.
XX
KW Soluble starch synthase; starch fine structure; corn;
KW transgenic plant; amylose; amylopectin; amylose polymerisation;
KW non-granule bound starch synthase; non-GBSSI; altered starch; food;
KW paper; plastic; adhesive; ss.
XX
OS Zea mays.
XX
PN WO200006755-A2.
XX
PD 10-FEB-2000.
XX
PF 26-JUL-1999; 99WO-US16296.
XX
PR 28-JUL-1998; 98US-0094436.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Broglie KE, Lightner JE;
XX
PS WPI; 2000-195311/17.
XX
PT Producing transgenic cereal crops with altered starch structure useful
PT for preparing foodstuff, paper, plastic or adhesives, comprises
PT transforming crops with chimeric sense or antisense gene construct
PT encoding starch synthase
XX
PS Claim 5; Page 51-52; 56pp; English.
XX
CC The present sequence is the corn soluble starch synthase (SSI) DNA
CC sequence inserted into plasmid pSS42. Chimeric gene of pSS42 was cloned
CC into the vector pKS17 for the generation of an antisense construct for
CC suppression of SSI expression in corn. The starch fine structure derived
CC from a grain of the cereal crop can be altered in the transformed cereal
CC crop by changes in amylose to amylopectin ratio, amylopectin fine
CC structure, increased abundance of very short amylopectin chains and in
CC the degree of polymerisation of amylose. These modifications can be
CC created by controlling the expression of non-GBSSI (non-granule bound
CC starch synthase) in transgenic plants. Altered starches are useful in
CC foods, paper, plastics or adhesives.
XX
SQ Sequence 1528 BP; 440 A; 383 C; 300 G; 405 T; 0 other;
```

```
alignment_scores:
  Quality: 2495.00      Length: 508
  Ratio: 5.051         Gaps: 1
Percent Similarity: 97.244 Percent Identity: 89.567
```

alignment_block:

US-09-674-824-2 x AA250637/rev ..
Align seg 1/1 to reverse of: AA250637 from: 1 to: 1528

```

134 GlnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPr 150
      ::::::::::::::::::::
1527 CGAGCTAAAGTAAACAACAAGCATTTGCTTTGTAACCGGGAAGCTTCCTC 1478
      ::::::::::::::::::::
150 oTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleA 167
      ::::::::::::::::::::
1477 TTATGCAAACTCTGGGGTCTAGGAGATGTTGTGGTTTCATTTGCCAGTTG 1428
      ::::::::::::::::::::
167 laLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeu 183
      ::::::::::::::::::::
1427 CTCCTTGCTGCTGCTGCTCACCGTGTGATGGTTGTAATGCCAGATATTTA 1378
      ::::::::::::::::::::
184 AsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHI 200
      ::::::::::::::::::::
1377 AATGGTACTCCGATAAGAAATTTATCAATGCATTTTACAGAAAAACA 1328
      ::::::::::::::::::::
200 stleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
      ::::::::::::::::::::
1327 CATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATG 1278
      ::::::::::::::::::::
217 luTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHis 233
      ::::::::::::::::::::
1277 AGTATAGAGATTACGTTGACTGGGTGTTGTTGATCATCCCTCATATCAC 1228
      ::::::::::::::::::::
234 ArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGI 250
      ::::::::::::::::::::
1227 AGACCTGGAAATTTATATGGAGATAAGTTGGTCTTTTGTGATAATCA 1178
      ::::::::::::::::::::
250 nPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleL 267
      ::::::::::::::::::::
1177 GTTCAGATACACACTCTCTTCTGCTGCTGATGTGAGGCTCCTTTGATCC 1128
      ::::::::::::::::::::
267 euGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsn 283
      ::::::::::::::::::::
1127 TTGAATTGGGAGGATATATTATGGACAGAAATTCATGTTTGTGTCAAT 1078
      ::::::::::::::::::::
284 AspTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPr 300
      ::::::::::::::::::::
1077 GATTGGCATGCCACTCTAGTGCCACTCTCTCTGCTGCAAAATATAGACC 1028
      ::::::::::::::::::::
300 oTyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuA 317
      ::::::::::::::::::::
1027 ATATGGTGTATTATAAGACTCCCGCAGCAVTTCTTGTATAATCAATTTAG 978
      ::::::::::::::::::::
317 laHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 333
      ::::::::::::::::::::
977 CACATCAGGGGTAGAGCCCTGCAAGCACATATCTTGACCTTGGGTGCCA 928
      ::::::::::::::::::::
334 ProGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgAr 350
      ::::::::::::::::::::
927 CCTGAATGGTATGGAGCTCTGGAGTGGGTATTCCCTGAATGGCGAGGAG 878
      ::::::::::::::::::::
350 qHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValV 367
      ::::::::::::::::::::
877 GCATGCCCTTTGACAGGGGTGAGCAGCTTAATTTTGAAGGTGCAGTTG 828
      ::::::::::::::::::::
367 alThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluVal 383
      ::::::::::::::::::::
827 TGACAGCAGATCGAATCGTGACTGTCAGTAAGGGTTATTTCGTGGGAGGC 778
      ::::::::::::::::::::
384 ThrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLy 400
      ::::::::::::::::::::
777 ACAACTGCTGAAGGTGACAGGGCCCTCAATGAGCTCTTAAGCTCCAGAAA 728
      ::::::::::::::::::::
400 sSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnP 417
      ::::::::::::::::::::
727 GAGTGTATTAAACGGAATTTGTAATGGAATTTGACATTAATGATTGAACC 678
      ::::::::::::::::::::

```

```

417 roThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSer 433
      ::::::::::::::::::::
677 CTGCCACAGACAAATGTATCCCTGTCTATTATTCTGTTGATGACCTCTCT 628
      ::::::::::::::::::::
434 GlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVa 450
      ::::::::::::::::::::
627 GGAAGGCCAAATGTAAAGGTGCATTCAGAAAGGAGCTGGGTTTACCTAT 578
      ::::::::::::::::::::
450 laArgLysAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnL 467
      ::::::::::::::::::::
577 AAGCGCTGATGTTCTCTGATGCTGCTTTATTTGGAAGGTTCGATTATCAGA 528
      ::::::::::::::::::::
467 ysGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAsp 483
      ::::::::::::::::::::
527 AAGCATGTGATCTCAATCACTTATCATACAGACTCTCATCGCGGAAGAT 478
      ::::::::::::::::::::
484 ValGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMe 500
      ::::::::::::::::::::
477 GTTCAATTGTTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGAT 428
      ::::::::::::::::::::
500 taArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyP 517
      ::::::::::::::::::::
427 GAGATCTACAGAGTCGATCTTCAAGGATAAATTTTCGTGGATGGTTGGAT 378
      ::::::::::::::::::::
517 heSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeu 533
      ::::::::::::::::::::
377 TTAGTGTTCAGTTTCCACCGAATAACTGCCGCTCGCATATATTGTTA 328
      ::::::::::::::::::::
534 MetProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGI 550
      ::::::::::::::::::::
327 ATGCCATCCAGATTCGAACTTCTGGTCTCAATCAGCTATATGCTATGCA 278
      ::::::::::::::::::::
550 nTyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrV 567
      ::::::::::::::::::::
277 GTATGGCACAGTTCTCTGTTCTCCATGCAACTGGGGGCTTAGAGATACCG 228
      ::::::::::::::::::::
567 alGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrp 583
      ::::::::::::::::::::
227 TGAGAGAACTTCAACCTTTTCGGTGAGAATGGAGAGCGGTACAGGGTGG 178
      ::::::::::::::::::::
584 AlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAl 600
      ::::::::::::::::::::
177 GCATTGCGACCCCTAACCCACAGAAAACATGTTGTGGACATTGCGAACTGC 128
      ::::::::::::::::::::
600 aMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysA 617
      ::::::::::::::::::::
127 AATATCTACATCAGGGACACAAAGTCCCTCTGGGAAGGGCTAATGAAGC 78
      ::::::::::::::::::::
617 rgGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSe 633
      ::::::::::::::::::::
77 GAGGCATGTCAAAAGACTTCACGTGGGACCATGCCCTGCAACAATACGAA 28
      ::::::::::::::::::::
633 rArgSerSerSerGlyProSer 640
      ::::::::::::::::::::
27 CAATCTTCCAGTGGGCCCTTCA 6
      ::::::::::::::::::::
seq_name: /SIDS1/gcgdata/geneseq/geneseq/NAL999.DAT:AA34652
seq_documentation_block:
ID AA34652 standard; DNA; 10336 BP.
XX
AC AA34652;
XX
DT 05-JUL-1999 (first entry)
XX
DE Wheat starch soluble synthase I (SSS I) gene sequence.
XX
KW Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;
KW starch branching enzyme; starch soluble synthase; debranching enzyme;
KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
KW grain softness protein I; bacterial isoamylase; glycogen synthase;
KW wSBE I-D4 gene; ds.
XX

```

OS XX Triticum tauschii.
FH FT Key Location/Qualifiers
FT FT CDS 1..9713
FT FT /tag= a
FT FT /product= "SSS I"
FT FT /note= "contains introns"
FT FT 1..316
FT FT /tag= b
FT FT /number= 1
FT FT 317..1471
FT FT /tag= c
FT FT /number= 1
FT FT 1472..1828
FT FT /tag= d
FT FT /number= 2
FT FT 1829..2765
FT FT /tag= e
FT FT /number= 2
FT FT 2766..2823
FT FT /tag= f
FT FT /number= 3
FT FT 2824..2905
FT FT /tag= g
FT FT /number= 3
FT FT 2906..3028
FT FT /tag= h
FT FT /number= 4
FT FT 3029..4112
FT FT /tag= i
FT FT /number= 4
FT FT 4113..4194
FT FT /tag= j
FT FT /number= 5
FT FT 4195..4285
FT FT /tag= k
FT FT /number= 5
FT FT 4286..4459
FT FT /tag= l
FT FT /number= 6
FT FT 4460..4561
FT FT /tag= m
FT FT /number= 6
FT FT 4562..4643
FT FT /tag= n
FT FT /number= 7
FT FT 4644..4743
FT FT /tag= o
FT FT /number= 7
FT FT 4744..4855
FT FT /tag= p
FT FT /number= 8
FT FT 4856..4998
FT FT /tag= q
FT FT /number= 8
FT FT 4999..5021
FT FT /tag= r
FT FT /number= 9
FT FT 5022..5101
FT FT /tag= s
FT FT /number= 9
FT FT 5102..5192
FT FT /tag= t
FT FT /number= 10
FT FT 5193..8592
FT FT /tag= u
FT FT /number= 10
FT FT 8593..8718
FT FT /tag= v
FT FT /number= 11
FT FT 8719..8806
FT FT /tag= w
FT FT /number= 11

FT exon 8807..8915
FT /tag= x
FT /number= 12
FT 8916..8991
FT /tag= y
FT /number= 12
FT 8992..9104
FT /tag= z
FT /number= 13
FT 9105..9160
FT /tag= aa
FT /number= 13
FT 9161..9199
FT /tag= ab
FT /number= 14
FT 9200..9497
FT /tag= ac
FT /number= 14
FT 9498..9713
FT /tag= ad
FT /number= 15
XX
PN WO9914314-Al.
XX
XX 25-MAR-1999.
PD
XX
XX 11-SEP-1998; 98WO-AU00743.
XX
PR 20-MAR-1998; 98AU-0002509.
PR 12-SEP-1997; 97AU-0009108.
XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD.
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Li Z, Morell M, Rahman S;
XX
XX WPI; 1999-229525/19.
DR P-PSDB; AAY09004.
XX
XX
PT New isolated cereal plant enzyme genes used for, e.g. expression of
PT antisense sequences of granule bound synthase
XX
PS Claim 10; Page 87-94; 171pp; English.
XX
CC The invention relates to a novel enzyme of starch biosynthetic pathway
CC in a cereal plant, where the enzyme is selected from starch branching
CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching
CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
CC SBE I of rice or maize. The methods and products can be used for
CC targeting expression specifically to the endosperm of the seeds of cereal
CC plants such as wheat or barley. They can be used for the expression of
CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
CC mol. wt. glutenin, grain softness protein I, bacterial isomylase,
CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
CC can be used for modifying the characteristics of starch produced by a
CC plant. The present sequence represents the wheat SSS I gene sequence.
XX
SQ Sequence 10336 BP; 2733 A; 2055 C; 2625 G; 2921 T; 2 other;

alignment_scores:
Quality: 2349.00 Length: 3317
Ratio: 3.149 Gaps: 25
Percent Similarity: 22.490 Percent Identity: 22.339

alignment_block:
US-09-674-824-2 x AAX34652 ..

Align seg 1/1 to: AAX34652 from: 1 to: 10336

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17


```

|||||
1  ATGGCGGCACGGGCTCGCGCGCGGGTGCCTCGCCCCCAGGCTCGGCT 50
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
51 GCGCGCGGATCCGGCGAGCGCGCGCGCGCGCTTGCCTGCGTCCGCGG 100
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
101 CGCGGCTCGGCGCTTGGCGGGGCGCGCTACGTCGCCGAGCTCAGCAGG 150
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67
151 GAGGCGCGCGCGCGCGCGCGCGCGCGAGCAGCAACTGGCCCGCGGCT 200
67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
201 CGTGCCAGGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 250
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
251 CCCCAGCAGCGCGCGCGCTGCGGAGCGCGCGCTGGGGAACTCGCGGCC 300
101 AspLeuLeuLeu Glu
|||||
301 GACCTCCTGCTCGAAGGTAAAAACAAGGCTGAATCCTCAGATCACCTCG 350
105
351 CGCTCTCGTTTTACAAATACGGTACTCGGAAGTGGTGTATATATGA 400
105
401 AGTTCTGTCGATTCTTCCTGACGGATGTTACGTCGATTGAGTTGATA 450
105
451 TATGTATACGTTGTTGTCATCGATCGTACAGATTTACCAGCACACTA 500
105
501 GATAGAAATCGAGACCGACGCGGCGAGATCAATAGATTTTCTAGAGTT 550
105
551 TTATTGGATCGTAGATGATTGATGGGGTGGCGGTGTCGATACGATAGCG 600
105
601 GTGCACCGCGGATGATATCGGGGATGTGCACGTGGTGGGTCTCAGCAGA 650
105
651 CATATCACTAGACTGGTATCGTAATTACTAGTACTGGAAGAGGAC 700
105
701 TAAAAAGGCTAGGCAAGTGCACGATGTTGGCAACGTTGTTAAATGAT 750
105
751 GAGTTTGCTCTTGGCTGGGCTGGTATTATTACCAAAAAAATGGTGTAGT 800
105
801 CCCTGTACTTATTAAATGGGAAATCTTAACATGACACTGGGGTTTAGAG 850
105
851 TCTCCAATTGTATATCTCAGCACTCAACTGATTTTACTGATACGTACT 900
105

```

```

901 GAAATGACACGTCGAGCACCCCTTCAAGGAATGCAATGCTTCTTCTG 950
105
951 TTTTATATTACAGGAACCTAGAAGAGAGTTCCACCTTTTGAGTACAGAA 1000
105
1001 CTCCTCCGTTCCAAAAATAGATGACTCAACTTGTGTACTAATTTGTACTA 1050
105
1051 TAGTTAGTACAAAGTTGAGTCATCTATTATTAGAACGGAGGAGTAGTATC 1100
105
1101 GAAATTGAAGACCCCTTGTTACTGCTTCTGTTTCAATCAAAATGGGAG 1150
105
1151 GCCCATGCAGTAAGTCACATGGGCACCTGGGAGGCTGGGATCATGTGTGC 1200
105
1201 TTTGCAGAGTACTAGACCCAGCTCACCCCTCTGTTAGATTACTTGTGGGC 1250
105
1251 TGCTACTTTTGTGTTGCTGTGCAGTATATCAGACATCCTGAATTTGGCAT 1300
105
1301 CTAGCTGAGAACAAGATGAGGTTGCACCATCTTATTATTGCTAAACTG 1350
105
1351 TTGTCACGCAATTTATAAAGAATGTGATCTTCTGAGTATTAAATTAATCAT 1400
105
1401 GTTCTGCTAATATCTGCTCCTGCTCTGGTGTGACAAATATACCATATGA 1450
106
1451 ATATTTTCCATTTTGCAACCAGGATGCTGAGGATTCATCGACAGCAT 1500
115
1501 AATCGTGGCTGCAAGTGAGCAGGATTCAGATCATGGATGCAATGAGC 1550
132
1551 AACCTCAAGCTAAAGTTACACGTAGCATCGTGTGTTGACTGGTGAAGCT 1600
149
1601 GCTCCTTATGCAAGTCAGGGGCTGGGAGATGTTGTTGCTGCTGTTACC 1650
165
1651 AATTCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
182
182 yLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAla 198
1701 ACTTGAATGGGTCTCTGTATAAAACTATGCAAGGCATTATACACTCGG 1750
199
1751 AAGCACATTAAAGATTCCATGCTTTGGGGATCATGAAGTGACCTTTT 1800
215
1801 TCATGAGTATAGAGACAACGTCGATTGGTGGGTACACAATCACCTTCTT 1850

```

```
224 ..... 224
1851 ATTCTGTGTAATTGAGCAACTGTTTATCCTTGTTACACTTCITTTA 1900
224 ..... 224
1901 GCCCTGCANAGACATATGTGATTTCCATACITTTTTTTGTTATTTCCCTGT 1950
224 ..... 224
1951 ACTCTTGCTCATGAAGTCAAAATATCATATATCATCCATGGAAGTCATCCAT 2000
224 ..... 224
2001 GTGCTAGTATTTTTGGTGCGGTGCCCTTTAACTTTTCAGGATTAATACG 2050
224 ..... 224
2051 TGGAAATTTGATAACTAAAGTTTATTTTATTGAAAAAAATTTAGGTTGGT 2100
224 ..... 224
2101 GAGCCACAGCCACGACGTGSCACCACCTGCTTGACACATGATTTTGCATTT 2150
224 ..... 224
2151 CTGTTTGACCGGACGACTTCATGTGAATAAGGTGTAAATCATAAAGTAC 2200
224 ..... 224
2201 CAATTTTATCTGCCAATTCGACTTAAGAGTATATACATTTATCTTGGCC 2250
224 ..... 224
2251 TCAATCATGGGAGTACTGTGCANTCAGTGCACCATCATTTGTTCTAAGGAG 2300
224 ..... 224
2301 AAAATGTGGTGCAAGGAAGACACTTTTGTCCCTTAATAAAGGCAGGCA 2350
224 ..... 224
2351 CTCCTGTGTGCATATAGATAGAAAGCAACAACCTTATTTCAAAGAGCTAAC 2400
224 ..... 224
2401 AATGCCAAAGAACCMAAANAAGCATGCTAAGCGGTGACACCAAAAGGT 2450
224 ..... 224
2451 GAGGGGGCCTTGTGACTGACAGCACCCCAAACTATTGCCATTGTTTAC 2500
224 ..... 224
2501 TAAATGAAGATCATTTTAGAAGCTCTCAGGAACCTTCGAAAACAGTGGCTT 2550
224 ..... 224
2551 TCCGTCACAGATCGTCTGTTAATATTTTGTCCAGTGATACTTTTTTTTG 2600
224 ..... 224
2601 CTCTTTACAAGAGTGCCTATGTTGACATATACATTTGTTAAGTTGTTTATA 2650
224 ..... 224
2651 AGTTTACTTCTTATTCTAAACAGCAAGTGCCTAAATGCTTGCATTTATTTT 2700
225 ..... Val. Phe 226
2701 GGCTATTTATTTTATTCATTTCATTCATCAACACTTTTGTTCAGGTGTTT 2750
```

```
227 ValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPh 243
|||||
2751 GTCGATCATCCGTCATATCATACACAGGAAGTTTATATGGAGATAATTT 2800
243 eGlyAlaPheGlyAspAsnGln. Phe. .... 251
|||||
2801 TGGTGCTTTTGGTGATAATCAGGTACACTACACTATACTAAGCTCCTAGT 2850
251 ..... 251
2851 TGACTAAGTCGTAAGTTGTACCTCCCTCGCTGACCGGCTGCTCTATGTCGT 2900
252 ..... ArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuI 266
|||||
2901 GCAGTTTCAGATACACACTCCTTTTGTCTATGCTGCATCGAGGCCCACTAA 2950
266 leLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValVal 282
|||||
2951 TCCTTGAATGGGAGGATATATTTATGGACAGAAATTCATGTTGTTGTTG 3000
283 AsnAspTrpHisAlaSerLeuValPro..Val. .... 292
|||||
3001 AACGATTTGGCATGCCAGCCTTGTGCCAGTGTACGTTGTTTGTGGATCTGA 3050
292 ..... 292
3051 AAGTCCAATCCTTTATTCATTCCTCTGCTTGTGCAGTGTGCCCATGCTCTACA 3100
292 ..... 292
3101 TTCTCTTTATGCTTTTTCATGCTGTTCTTATATATTGCATATATGCTTAT 3150
292 ..... 292
3151 GGAGTCTAAAAGTTACCGGAGGAATAACTCTTAAGGATTTTCCTCAATCA 3200
292 ..... 292
3201 ATTATCTTTAGCTTTTATTAGATTAACTTTACTGTGGCAACATAATGTGTTTT 3250
292 ..... 292
3251 GAGATTTACAAGTTTCAGAGATTGCACCTTCACCTAGTTGCTAGCTAATCTGA 3300
292 ..... 292
3301 TGTTTTCCCGAGAAAATGCCTAAAGCTTTGTCTTTGATGCTGATGATAG 3350
292 ..... 292
3351 AAAAAGAGTTTATGTACACTCCCAAGAGGGGACCCCAAAATTAACAACACC 3400
292 ..... 292
3401 ACACCCCTGAGAAGTCTAGGCGCTGCCGGAAGAGCGATGCAAGCCCCACTG 3450
292 ..... 292
3451 CCCCTGCTTAGCTCAAGCCGGCGTCAGCTTTCATTTGTGTCAAGTAAGC 3500
292 ..... 292
3501 TAGCAGTCTAGATTGCGCAAGGTCGATTCGTCGAAGATGACAGTGTTCG 3550
292 ..... 292
3551 GCTGCTTCCAAATCCACCAAACTATGACATGATCAGTGGAGAAAGTACCT 3600
292 ..... 292
3601 TTTCTCGCGGTGAGGGGTGAGCTGGTGGTCTGCTGCTGCCAGTTTTC 3650
292 ..... 292
```

3651 GATAATCTGAAAAATGCATGTTTTCATGATTTTAGTATCTTTGCGGACCT 3700
292 292
3701 GGGTACCACCTAAGCTTTTCACACAGATAATTTGCAGTTTACACCTATAAAG 3750
292 292
3751 TAACGGTCATGATATGCATGTGTTTGGGTAGATCATGGTGCATGCATTT 3800
292 292
3801 TAGGAATTAGGACATGCCAGAACACAGCTGAGGCTTATGGGCAATTCAAT 3850
292 292
3851 TGTTCATATACGAGTCATCAATATGGTTCAGCATGTTTGGACGCTACT 3900
292 292
3901 TGTTTGGGCAATTTCAGATGGTGAATGTAGCTGCTTGATGTTGGCTAG 3950
292 292
3951 CTGCTTATTTGTACAAAGTATCGATGTTAGATGCATATTTCCCTTTTGT 4000
292 292
4001 CTTGTGCTGTTTGGCATGTTGTATTTCCCTTTTCTGTGCGCAGTTGCA 4050
292 292
4051 TGTAAATTTGTTTTCATTAATACATAATCAACTTTTGTGCTGACATCAGTCA 4100
293 LeuLeuAlaLysTyrArgProTyrGlyValTyrAr 305
4101 TTTTATTACGCTTCTGTGCAAAATATAGACCATACCGGTGTTACAG 4150
305 gAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGln..... 319
4151 AGATTCCCGCAGCACCTTGTATACATAAATTTAGCACATCAGGTTTGGG 4200
319 319
4201 TCTATCACCTTTTCATTAATCCGTACATGGCTTTGTAAAGTCGGTTCACACGT 4250
320 Gly.ValclupProAlas 325
4251 ATCGTCATCTGTATGTTATTTCATGTCATTAGGGTGTGGAGCCTGCCAA 4300
325 erThrTyrProAspLeuGlyLeuProGluTyrTrpTyrGlyAlaLeuGlu 341
4301 GTACATATCTGATCTGGGATTTGCCACCTGAATGGTATGGAGCTTTAGAA 4350
342 TrpValPheProGluTrpAlaAlaArgHisAlaLeuAspLysGlyGluAl 358
4351 TGGGTATTTCCAAATGGGCAAGAGGATGCCCCCTTGACAAGGGGTGAGGC 4400
358 aValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrV 375
4401 AGTTAACTTTTGAAGGAGCAGTTGTGACAGCAGATCGAATTTGTGACCG 4450
375 alSerGln..... 377
4451 TCAGTCAGGTGAATACTCAATACTTCTCTTTTCTTTTTCGCGGATGTTTC 4500
377 377
4501 TTCAGTTCAATTCCTGCTGCTTTCACCCCAATTAAGAAATGATTTAATCTT 4550
378 GlyTyrSerTrpGluValThrThrAlaGluGlyGlyGlnG 391
|||||

4551 TTGTTTCTAGGGTTATTTCATGGGAGGTCACAACCTGCTGAAGGTGGACAGG 4600
391 lyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGly.Ile... 406
|||||
4601 GCCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGTATTGAATGGTAACTAT 4650
406 406
4651 ATTTGAATCCACTTATCTTCTTCTGAACATATTTTACAGAAATAGATGGA 4700
407 Val 407
4701 TGGGTTCCAGAATAAATTTCAGTTTGGCTTTCGGGTATGAAGGAATGTGA 4750
408 AsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuPr 424
|||||
4751 AATGGAATTGCATTAATGATTGGAACCCCAACACAGACAGCAAGTGTCTCCC 4800
424 oHisHisTyrSerValAspAspLeuSerGlyLys..... 435
|||||
4801 TCATCATTTATCTGTCATGACCTCTCTGGAAAGGTGTGTGGATAGTACC 4850
435 435
4851 CTATATAATAACATGATATATCTGATCTAGTACTTCTTTTCTTTGCTAG 4900
435 435
4901 TTTGCTTCCCATGATGTTTCTCACTAACTAATCTATGTTGTTTGGCATAC 4950
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProV 450
|||||
4951 TTGTCAGGCCCAATGTAAAGCTGAATTCAGAGAAGGAGTGGGTTTACCTG 5000
450 alArgGluAspValProLeu..... 456
|||||
5001 TAAGGAGGATGTTCTCTCTGTTAGATACAAACCCCTTAAGATATATATT 5050
456 456
5051 TTTAAATCCCTAAAAAACTTGCCGATCATCTCATTTAGCTTGTATTCACA 5100
457 lLeGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleL 473
|||||
5101 GATTGCTTTTATTGGAAGACTGGATTACCAGAAAGCATTTGATCTCATTA 5150
473 ysMetAlaIleProGluLeuMetArgGluAspValGlnPhe..... 486
|||||
5151 AAATGGCCATTCCAGAGCTCATGAGGGAGGAGTTCAGTTTGTAAAGTTCA 5200
486 486
5201 TATTCTTTTCTTGAGACTAGAGTATAAATCAACATGTAGGTGGGGT 5250
486 486
5251 GGTATAATACAGACATAAGTTCACGCTATTGCTTCCATGAGAATTTAAT 5300
486 486
5301 GCTATTTCAGTAATATGCTACTGCAAGTTTGAACAAAAAGTTGGAAGCAAT 5350
486 486
5351 AAATATATGTGTAGCAGCTGACCATGCAGTGCACCTATAGCTGGAATGTCC 5400
486 486
5401 TGAGTCTATGTGATCTAAACACACTCAACACATGTTTTTTCGCATACAAAC 5450
486 486
5451 ACATCGGTGGCGCAACAAACATACTCTACATAAAAAATTTGGCTTGGTGA 5500

486 486
5501 CTGCAGACATGCTCTTATCTCCATTCCAACATTTCTTGTTCACATGG 5550
486 486
5551 CTGAAGACTAAGAAAGGGGGACCCAGGGTGATGTAGCCAACTAGATCCA 5600
486 486
5601 GTAAGGAAGCTAGCCGAGCCTAGGAGATTGCTTAGGTAGCTGGAACGT 5650
486 486
5651 AGGCTCTGACAGGGGAAGCTTCCGGGAGCTAGTCGATGCCAGTGGTGAGGA 5700
486 486
5701 GAGGTGTGATATCTTTTGCCTCCAAGAAACCAAAATGTAGGGGACAGAAG 5750
486 486
5751 GCGAAGGAGGTGGAGGATACCGGCTTCAAGCTGTGGTACATGGGACGGCT 5800
486 486
5801 GCAACAGAAATGGCGTAGGCATCTTGATCAACAAGAGCCCTTAAGTATGG 5850
486 486
5851 AGTGTAGACGTCAGAGACGTGGGGACCGGATTATCCTCTCAAGCTGG 5900
486 486
5901 TAGTTGGGACTTAGTCTCAATGTTATCAGCGTGTATGCCCGCAAGTA 5950
486 486
5951 GCCCACAATGAGAACGCCAAGAGGAGTTCCTGGGAAGCCCTGGAAGACAT 6000
486 486
6001 GGTAGGAGTGTACCGATTGGCGAGAAGCTCTTCATAGAGGAGACCTCA 6050
486 486
6051 ATGGCCACGTGGGTACATCACTAATAGGTTTGAAGGGGCACATGGGGC 6100
486 486
6101 TTTGGCTATGGCATCAAGAATCAAGAAGAAGATGTCTTACGCTTTGCTCT 6150
486 486
6151 AGCCTAGCACATGTTGTAGCTAACACCCCTCTTTAGAAAGAGAGATCAC 6200
486 486
6201 ATCTGTGTACTTTTAGTAGTGGCCAACTAGCCAGATCGATTTTCATGCT 6250
486 486
6251 CTCGAGAAGAAGATAGGTGTCGCGCCTAGACTGCAAGGTGATACCTTT 6300
486 486
6301 CGGATTCTGTCCAGCGGGATTAAGCCTGCCAAAGTCGCTAGATGAAGTG 6350
486 486
6351 GTGGAAGCTCAAGGGGGAGGTAGCTCAGGCGTTCAAGGAGAGGGTCAATTA 6400

486 486
6401 GGGAGGCCCTTGGAGGAAGGAGGGGATCCGACAATGTGTGATGAAG 6450
486 486
6451 ATGGCACTTGCAATTCGTAAAGTGGCCTCGGAGGAGTGTGGAGTGTCCAG 6500
486 486
6501 GGATGGAGNAGCAAGATAAGGATACCTGGTGGTGAATGATGATGTCC 6550
486 486
6551 AGAAGGCAATTAAAGAGAGAAAGATTGCTTTAGACGCCCTATACTTTGGAT 6600
486 486
6601 AGGAGTGCAGTCAACATAGAAAAGTACAAGATGGCGAAGAGCCCCANA 6650
486 486
6651 GCGAGCTGTCAAGTGAAGCAAGGGTGGGCATATGAGGATCTCTACCAAC 6700
486 486
6701 GGTAGGCACNAAGAGAGCGGATATTGGCCCAAGTCAATAGATGGCCAAATC 6750
486 486
6751 CGAGAGAGAGAGACAGCGGATATTGGCCCAAGTCAATGCAATCAAGGAT 6800
486 486
6801 GGACAGACCAACTCTTGGTGAAGGACGAGGAGATTAAAGCATAGATGGCG 6850
486 486
6851 GGATACTTCGACAAGCTGTTCAATGGGGAGGATGAGAGTCTTACCATTG 6900
486 486
6901 AACTTGACGACTCCTTTGATGAGACCATCATGCGTTTATCGCGCAATC 6950
486 486
6951 CAGGAGTCCGAGGTCAAGGAGGCTTTAAAAGGAGCAAGCGATGGGCC 7000
486 486
7001 CTGATTGTATCCCATTCAGGTGTGAAAGGCCCTCGGGGACATAGCGATA 7050
486 486
7051 GTATGGCTAACCAAGCTATTCAACCTCATTTTCGGGCAACAAGATGCC 7100
486 486
7101 AGAAGATGGAGAGAGATATATTAGTACCATCATCAACAGGGGGGAT 7150
486 486
7151 GTTCAGAGTTGTACTAATTACCATGGAATTAAAGCTGATGAGCCATACAAT 7200
486 486
7201 GAAGCTATGGAGAGAAATCATTTGAGACCCGCTTAAGAAGATGACAAGCG 7250
486 486
7251 TGACCAAAAATCAGTTTGGTTTCATGCTTGGAGGTCGACCATGGAACC 7300
486 486

7301	ATTTTCTTGGTACGACAACCTTATGGAGAGATACAGGAGCAAAAGAAGGA	7350
486	486
7351	CTTGCATATGGTTCATTGACTTGNAGAAGCCCTATAATAAGATACCGC	7400
486	486
7401	GGAAATGTCATGTGGTGGCCCTTGGAGAAACACAAAGTCCAGCAAAAGTAC	7450
486	486
7451	ATTACCCTCATCAAGGACATGTACCATAATGTTGTGACAAGTGTTCGAAC	7500
486	486
7501	AAGTATGTCGACACTAATGACTTCCCGATTAAAGATAGGACTGCATCAGG	7550
486	486
7551	GGTCAGCTTTGAGCCCTTATCTTTTGGCTTGGTGATGGATGAGTCA	7600
486	486
7601	AGGATATACAAGGAGATATCCATGCTATGCTCTTTGTGATGATTT	7650
486	486
7651	GGTCTAGTTGACGATAGTCGGCGGGGTAAATAACAAGTTAGAGTTAT	7700
486	486
7701	GGACACAAACCTTGAATCGAAAGGTTTAGGCTTAGTAGAACTAAAACC	7750
486	486
7751	GAGTACATGATGTCGGTTTCAGTACTACTAGGTGTGAGGAGGAGGT	7800
486	486
7801	TAGCCTTGATGGCAGGTGTACCCAGAGGACACCTTTCGATATTGG	7850
486	486
7851	GGTCAATGCTGCAGGAGATGGGGTATTGATGAAGATGTGAACCATCGA	7900
486	486
7901	ATCAAGCTCGATGGATGAAGTGGCGCCAAGCTTCTGGCATCTTTGTGA	7950
486	486
7951	CAAGAGAGTGCCACAAAAAGCTAAGGCAAGTTCACAGGACGGCGTTTCA	8000
486	486
8001	CCCGCAATGTTGATGGCGCTGAGTGTGCCGACTAAAGGGGACATGT	8050
486	486
8051	TCAACAGTTAGGTGTGGCGGAGATGCGTATGTTGAGATGGATGTGTGCC	8100
486	486
8101	ACAGGAGGAAGTATCGAGTCCGGAATGATGATATACGAGATAGAGTTGG	8150
486	486
8151	GTAGCACCATTGAAGAGAGCTTGTCCAACATCGTCTGAGATGGTTGG	8200
486	486

[illegible]

```

595 ..... 595
9151 AAGTTTTTGGTGGACTCTTTGTCCGGTTATAGGATCGACCTTGGCTGTAGC 9200
595 ..... 595
9201 ATGGTACCTTAGTGCCCTTTGTATATAGACCTAAACGTATGAGCTACATT 9250
595 ..... 595
9251 TGTCTACACTAATCATAGTAGTCGATTGCCGGAGCGTTTTTGCCTGGAT 9300
595 ..... 595
9301 TCTGCTAATTTAATTTTCATGACGATAACTCATACCATGTTTGGTTCTC 9350
595 ..... 595
9351 CGATGGGGCCAGAAATGGCGTCTAGTGTCTGCCATCTGTCTAACTAGCCA 9400
596 ..... Ala... 596
9401 ATGCGGGTGTTCCTCAAGTGAAAAATTTACCTTTTGACCACTTGTCAGGCA 9450
597 LeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGluG 613
9451 TTGGCAACCGCGATGTCGACATTCCAGGGAGCACAAGCCGTCTCTGGAGGG 9500
613 yLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla.Pro 629
9501 GCTCATGAAGCGAGGCATGACGAAGAAGACCATACGTGGGACCATCGCCCG 9550
630 SerSerThrSerArgSerSerSerGlyProSerTrpThrAsnProThrSe 646
9551 AGCAGTACGAGCAGATCTTCGAATGGGCTTCGTGGACCAACCTCAGTC 9600
646 rCysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaLeuL 663
9601 ATGTAGCGGGACTGGGAGGTGCGAAGCGGGGTCTCTTGAGCTCTGA 9650
663 ysthrSerSerSerSerPheargGlyProGluGlyTyrProCysThrLeu 679
9651 AGACATGTTTCCTCATCTTCGCGCGCCGGAAGGATACCCCTGTACATTG 9700
680 ArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAl 696
9701 CGTTGTCTGCTACAGTAGAGTCGCAATGCCCTGCTTGTCT.TGGTCCGC 9749
696 aGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerG 713
9750 CGGTTCCAGATAGATAGCGGTCTGCTGCTCGCGCGGTGACAGCTTCGG 9799
713 lGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAlaAlaGly 729
9800 GTGGATGCAGGTTACAGTTTTTGGGCAATAGAGAGGGATGTGCTGCAGGA 9849
730 TrpLeuThrAlaLysHisSerAspGlySerLeuSerValArgValTh 746
9850 TGGTTAACAGCAAAACCACTCAGATGCAGCCCTCTCTGTCGCTGTTAC 9899
746 rAlaGluIleArgAsnGlnLeuValThrLeu 756
9900 AGCTGAATACAGAAACCACTGGTGACTCTT 9930
seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AA250642
seq_documentation_block:
ID AA250642 standard; cDNA; 1415 BP.
XX
AC
AA250642:
XX
TT 23-MAY-2000 (first entry)

```

```

XX      Corn soluble starch synthase gene fragment inserted in pSS64-C5.
XX
XX      Soluble starch synthase; starch fine structure; corn;
XX      transgenic plant; amylose; amylopectin; amylose polymerisation;
KW      non-granule bound starch synthase; non-GBSSI; altered starch; food;
KW      paper; plastic; adhesive; ss.
XX
XX      Zea mays.
XX
XX      WO200006755-A2.
XX
XX      10-FEB-2000.
XX
XX      26-JUL-1999; 99WO-US16296.
XX
XX      28-JUL-1998; 98US-0094436.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Broglie KE, Lightner JE;
XX
XX      WPI; 2000-195311/17.
XX
XX      Producing transgenic cereal crops with altered starch structure useful
PT      for preparing foodstuff, paper, plastic or adhesives, comprises
PT      transforming crops with chimeric sense or antisense gene construct
PT      encoding starch synthase
XX
XX      Claim 5; Page 52-53; 56pp; English.
XX
XX      The present sequence is the corn soluble starch synthase (SSI) DNA
CC      sequence comprising the SSI coding region of amino acids 1-494 inserted
CC      into plasmid pSS64-C5. The chimeric gene containing the zein promoter
CC      followed by the SSI gene fragment is used as a sense construct for
CC      preparation of transgenic corn expressing altered starch structure. The
CC      starch fine structure derived from a grain of the cereal crop can be
CC      altered in the transformed cereal crop by changes in amylose to
CC      amylopectin ratio, amylopectin fine structure, increased abundance of
CC      very short amylopectin chains and in the degree of polymerisation of
CC      amylose. These modifications can be created by controlling the
CC      expression of non-GBSSI (non-granule bound starch synthase) in
CC      transgenic plants. Altered starches are useful in foods, paper, plastics
CC      or adhesives.
XX
XX      Sequence 1415 BP; 317 A; 335 C; 397 G; 366 T; 0 other;
XX
XX
XX
XX
XX      alignment_scores:
XX      Quality: 2036.00      Length: 478
XX      Ratio: 4.735      Gaps: 4
XX      Percent Similarity: 89.958      Percent Identity: 81.799
XX
XX      alignment_block:
XX      US-09-674-824-2 x AAZ50642 ..
XX
XX      Align seg 1/1 to: AAZ50642 from: 1 to: 1415
XX
XX      4 ThrGlyValcylAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
XX      ::::|||||:|||||:
XX      14 TGGCCGCTGGGCGCGCGCTCTCTC..... 40
XX
XX      20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
XX      ||| ||||| :|||
XX      41 .CTCGCGCGCGCGCTGTGCGCGCGCGCTGCGGACCGCGCGCGCGCGC 89
XX
XX      35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
XX      ||| :|||:|||||:
XX      90 GCGGCTCTCAGCGCGTGTGCGCGCGCGCGTGTGCGCGAGCTGAGCAGG 139
XX
XX      51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProLeu 67
XX      ||||| |||||
XX      140 GAGGGCGCGCGCGCGCGCTGCGCACCGCGCTGCTGCGCGCGCGCT 189

```

67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
 190 CGTCCCGGGTTCCTCCG...CGCCGCGGAGCCACCGGTGAGCCGG 236
 84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
 237 CATCGACCGCCGCGCGTCCGACGCGCGCTGGGGACCTCGGT... 283
 101 AspLeuLeuGluGlyIleAlaGluAspSerIleaspSerIleIleVa 117
 284 ...CTCGAACCTGAAGGGATGTGAAGTTCATCGATAACACAGTAGT 330
 117 lAlaIAserGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
 331 TGTGGCAAGTGAGCAAGATTCTGAGATTGGTGGAAAGAGCAAGCTC 380
 134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
 381 GAGCTAAGTAACACAAAGCATTCTCTTTGTAACCGCGGAAGCTTCCT 430
 151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
 431 TATGCAAAAGTCTGGGGTCTAGGAGATGTTTGTGGTTTCATTGCCAGTGC 480
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
 481 TCITGCTGCTGGGTCCACCGTGATGGTGGTGGTGGTGGTGGTGGTGGT 530
 184 sNGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
 531 ATGCTACCTCCGATAAGAATTATGCAAAATGCATTTTACACAGAAACAC 580
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
 581 ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATGA 630
 217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHis 234
 631 GTATAGAGATTTCAGTTACCTGGGTGTTTGTGATCATCCCTCATATACA 680
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
 681 GACCTGGAAATTTATATGGAGATAAGTTTGGTGCTTTTGGTGAATAACAG 730
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
 731 TTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGATCCT 780
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
 781 TGAATTGGGAGGATATATTTATGACAGAGATTGCATGTTGTGTCATAG 830
 284 spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300
 831 ATTGCCATGCCAGTCTAGTGCAGCTCCTTCTGTGTCGCAAAATATAGACCA 880
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
 881 TATGGTGTATTAAGAAGACTCCCGCAGCATCTTGTATAACATAATTTAGC 930
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
 931 ACATCAGGGTGTAGAGCCTGCAAGCACAATATCCTGACCTGGGTGGCCAC 980
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
 981 CTGAATGGTATGAGCTCTGAGTGGGTATTCCCTGAATGGCGGAGGAGG 1030
 351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
 1031 CATGCCCTTGACAAAGGTGAGGAGCTTAATTTTGTGAAGGTGCAGTTGT 1080

367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
 1081 GACAGCAGATCGAATCGTGAATCGTGAATAGGGTTATTCGTGGGAGGTCA 1130
 384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
 1131 CAACTGCTGAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1180
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
 1181 AGTGTATTAAACGGAATTGTAATGGAAATTGACATTAAATGATGGAACCC 1230
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
 1231 TGGCACAGACAATGATCCCTGTCATTATTCGTGATGACCTCTCTG 1280
 434 lYlsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
 1281 GAAAGGCCAAATGTAAAGGTGCATTGCAGAGGAGCTGGGTTTACCTATA 1330
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
 1331 AGCCTGATGTTCTCTGATGGCTTTATTGGAAGTTGGATTATCAGAA 1380
 467 sGlyIleAspLeuIleLysMetalIleProGlu 478
 1381 AGCATTTGATCTCATTCACACTTATCATACCAGAT 1414

OM of: US-09-674-824-2 to: Issued_Patents_NA:* out_format : pfs
Date: Mar 28, 2002 5:43 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09674824/runat_27032002_094251_18439/app_query.fasta_1.824
-DB=Issued_Patents_NA -QPM=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DEGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MTNLEN=0 -MAXLEN=2000000000
-USER=US09674824 -CGN1_1_94 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-674-824-2
Query length: 756
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 113238999
Search time (sec): 92.300000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-196-390-1 +	3612.00	5385.06	2.7e-292	2239	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-192-909-1 +	2809.50	4180.73	3.2e-225	2383	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-12 +	2713.50	4040.48	2.1e-217	1752	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-572-951-1 +	2651.00	3940.26	8.0e-212	2990	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-9 +	2001.00	2968.27	1.1e-157	2360	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-3 +	1731.50	2567.65	2.3e-135	1758	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-10 +	1246.00	1837.36	1.1e-94	2097	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-7 +	1215.00	1787.39	6.6e-92	2793	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-5 +	1184.00	1745.41	1.4e-89	1926	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-196-390-5 +	1182.50	1738.51	3.5e-89	2825	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-8 +	1076.00	1582.93	1.6e-80	2007	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-572-931-2 +	1071.00	1574.97	4.5e-80	2085	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-07-735-065-1 +	943.50	1382.14	2.5e-69	2380	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-6 +	900.00	1316.10	1.2e-65	2542	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-735-065-1 +	678.50	990.59	1.6e-47	1464	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-469-202-11 +	678.50	990.59	1.6e-47	1464	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-735-491-1 +	634.50	923.52	8.6e-44	1601	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-470-720-2 +	612.50	884.88	1.2e-41	2549	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-470-720-5 +	604.50	864.80	1.6e-40	4964	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-4 +	603.00	862.96	2.0e-40	4800	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-115-704-1 +	558.50	803.41	4.2e-37	2652	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-08-836-567-1 +	556.00	801.37	5.5e-37	2303	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-11 +	556.00	794.18	1.4e-36	4168	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-968-542C-1 +	526.00	744.71	7.9e-34	6027	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-370-700-1 +	157.50	190.48	0.0058	6854	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-370-700-1 +	153.00	153.91	0.6355	80161	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-08-804-227C-7 +	152.00	159.58	0.3071	44377	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-804-198-1 +	152.00	159.58	0.3071	44377	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-07-755-573C-7 +	148.50	193.13	0.0042	1810	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-428-517-1 +	146.00	148.91	1.21	50937	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-642-734C-3 +	142.50	154.86	0.5628	20235	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-439-009A-3 +	142.50	154.86	0.5628	20235	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-398-291A-3 +	140.50	168.76	0.0946	5020	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-804-227C-7 +	140.50	142.33	2.80	44377	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-804-198-1 -	140.50	142.33	2.80	44377	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-443-501A-2 +	140.00	135.72	6.55	71989	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-097-997A-12 +	137.00	167.68	0.1087	3561	

/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-665-574C-12 + 137.00 167.68 0.1087 3561
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-08-946-994-12 + 137.00 167.68 0.1087 3561
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-614-770A-1 - 137.00 135.82 6.47 49272
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-658-136-2 + 137.00 134.81 7.36 53526
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-658-136-1 + 137.00 134.80 7.37 53577
seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-196-390-1
seq_documentation_block:
: Sequence 1, Application US/09196390
: Patent No. 6307125
: GENERAL INFORMATION:
: APPLICANT: Block, Martina
: APPLICANT: Lorz, Horst
: APPLICANT: Lutticke, Stephanie
: APPLICANT: Walter, Lennart
: APPLICANT: Froberg, Claus
: APPLICANT: Kossmann, Jens
: APPLICANT: Kossmann, Jens
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
: TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/196,390
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE 196 21 588.9
: FILING DATE: 29-MAY-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE 196 36 917.7
: FILING DATE: 11-SEP-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP97/02793
: FILING DATE: 28-MAY-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Haley, Jr., James F.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: AGREVO-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 596-9000
: TELEFAX: (212) 596-9090
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2239 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Triticum aestivum L.
: STRAIN: cv. Florida
: HAPLOTYPE: ca. 21 d Caryopses
: IMMEDIATE SOURCE:
: LIBRARY: cDNA library in pBluescript sk (-)
: CLONE: TASSS
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..2017

Fri Mar 29 09:26:44 2002

US-09-196-390-1

alignment_scores:
Quality: 3612.00 Length: 671
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-674-824-2 x US-09-196-390-1 ..

Align seg 1/1 to: US-09-196-390-1 from: 1 to: 2239

86 ThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102
3 ACGCAGCGCCCTGCGGACGCCGGCTGGGGAACCTCGCGCCGACCT 52
102 uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleValAlaA 119
53 CTGCTCGAAGGGATTGCTGAGGATTCATCGACAGCATAAATTGTGGCTG 102
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAla 135
103 CAAGTGACAGGATTCAGATCATGATGCGAATGAGCAACCTCAAGCT 152
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152
153 AAGATTACAGTAGCATCGTCTTTGTGACTGGTGAAGCTGCTCTTATGC 202
152 alySerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuA 169
203 AAGTCAGGGGGTGGGAGATGTTGTGGTTCGTTCACCAATGCTCTTG 252
169 laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185
253 CTCTCGTGTGCCAGGATGATGGTGTAAATGCCAAGATACTTAAATGGG 302
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLy 202
303 TCCTCTGATAAAACATATCAAGGGATTTATACACTGCGAAGCACATTAA 352
202 sileProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrA 219
353 GATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGAGTATA 402
219 rgAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 235
403 GAGACAACGTCGATTGGGTGTTTGTGTCATCCGTCATATCACAGACCA 452
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252
453 GGAAGTTTATATCGAGATAATTTTGGTGTCTTTTGGTGATAATCAGTTCA 502
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269
503 ATACACACTCCTTTGCTATGCTGATCGAGGCCGCCACCTAACTCTTGAAT 552
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrp 285
553 TGGGAGGATATATTTATGGACAGNAATTCATGCTTTGTTGTGAACGATTGG 602
286 HisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrG1 302
603 CATGCCAGCCTTGTGCCAGTCTCTTGTGCTGCAAAATATAGACCATACGG 652
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
653 TGTTTACAGATTCCCGACACCCCTTGTATACATAAATTTAGCACATC 702
319 lnglyValGluProAlaSerThrTyrProAspLeuGlyLeuProProGlu 335
703 AGGGTGTGAGCCCTGCAAGTACATATCTGTGCTGGGATTTGCCCTCTGAA 752

336 TrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAl 352
753 TGGTATGGAGCTTTAGAAATGGGATTTCCAGAAATGGCAAGGAGCATGC 802
352 aLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrA 369
803 CCTTGACAAGGGTGGAGGAGTTAACTTTTGAAGGAGGAGCATGTGTGACAG 852
369 laAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThr 385
853 CAGATCGGATTGTGACCTGACCTGACCTGAGGTTATTTCATGGAGGTCAAACT 902
386 AlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402
903 GCTGAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGT 952
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrT 419
953 ATTGAATGGAATTTGAAATGGAAATGACATTAATGATTTGGAACCCACCA 1002
419 hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435
1003 CAGACAAGTGTCTCCCTCATCTATTCTCTCGATGACCTCTCTGGAAG 1052
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgG1 452
1053 GCCAAATGTAAGCTGAATTCAGAAAGGAGTTGGGTTTACCTGTAAAGGA 1102
452 uAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469
1103 GGATGTTCTCTGATGGCTTTATTGGAAGACTGGATTACCAAGAAAGGCA 1152
469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485
1153 TTGATCTCATTAATAATGGCCATTCACAGACTCATGAGGAGGACGTGCRA 1202
486 pheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgse 502
1203 TTTGTCATGCTTGGATCTGGGGATCCAATTTTCAAGGCTGGATGAGATC 1252
502 rThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerV 519
1253 TACCGAGTCGAGTTACAAGGATAAATTCCTGGATGGTTCGATTTAGTG 1302
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535
1303 TTCCAGTTTCCACAGAAATAACTGCAGGTTCCGATATATTTGTTAATGCCA 1352
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG1 552
1353 TCGAGATTTGAACCTTGGGCTTTAATCAGCTATATGCTATGCAATATGG 1402
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569
1403 TACAGTTCTGTAGTTTCATGGAACCTGGGGGCTCCGAGACACACATCGAGA 1452
569 hrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPhe 585
1453 CCTTCAACCCCTTTGGTGCACAAAGAGAGAGGTTACAGGTTGGGCGGCTTC 1502
586 SerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSe 602
1503 TCACCGCTAACCGTGGACAAGATGTTGTGGCATTTCCGAACCGGATGTC 1552
602 rThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyM 619
1553 GACATTCAGGAGACACACCCCTCTCTGGAGGGGCTCATGAAGCGAGCA 1602
619 etThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSer 635
1603 TGACGAAAGACCATACGTGGGACCATGCCCCGAGCAGTACGAGCAGATCT 1652
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG1 652

```

|||||
1653 TCGAGTGGCCCTCGTGGACCAACCCCTACGTCATGTAGCGGGACCTGGG 1702
652 YArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerP 669
|||||
1703 GAGTCCAAAGTCCAGCTCTCTCAGCTCTGAAGACATCTCTTCATCT 1752
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685
|||||
1753 TCCGCGCCCGAAGGATACCCCTGTACATTCGCTTCTCTGTACAGTA 1802
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTrpAs 702
|||||
1803 GAGTCGCAATCGCGCTCTCTTGTGTCGCCGTTTCGAGAACATATGA 1852
702 pGlyCysAlaAlaAlaValThrAlaSerGlyArgGlnLeuGlnP 719
|||||
1853 CGGCTGTGCTGCTCGCGCGGTGACAGTTCGGGTGGACAGTTACAGT 1902
719 heTrpGlyLeuArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
|||||
1903 TTTGGGAATAAGGAGGATGTCTGCAGGATGTTAACAGCAAGCAC 1952
736 HisSerAspGlySerLeuSerValArgValThrAlaGluLeuArgAsnG 752
|||||
1953 CACTCAGATGCGACCTCTCTGTCGCTGTACAGCTGAATCAGAAACCA 2002
752 nLeuValThrLeu 756
|||||
2003 ACTGGTGACTCTT 2015

```

seq_name: /cqn2_6/ptodata/1/1na/6B_COMB.seq:us-09-192-909-1

```

seq_documentation_block:
; Sequence 1, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Frohberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02527
; FILING DATE: 16-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 2383 base pairs
; TYPE: nucleotide
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; TISSUE TYPE: endosperm
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1950
; OTHER INFORMATION: /function= "starch synthesis"
; OTHER INFORMATION: /product= "soluble starch synthase"
US-09-192-909-1

```

```

alignment_scores:
Quality: 2809.50      Length: 749
Ratio: 4.411          Gaps: 10
Percent Similarity: 85.047      Percent Identity: 73.565

alignment_block:
US-09-674-824-2 x US-09-192-909-1 ...
Align seg 1/1 to: US-09-192-909-1 from: 1 to: 2383

```

```

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
|||||
41 TCGGCGGTGGCGCGCGCTGCTCTCCTC..... 67
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
|||||
68 CTGCGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCGCG 116
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
117 GCGCGCTCCAGCGCGTGTGCGCGCGCGTGTGCGCGAGGTGACGAG 166
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
|||||
167 GAGGCGCGCGCGCGCGCGCTGCTGCGCGCTGCTGCGCGCGCGCT 216
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
|||||
217 CGTGCGCGCTTCTCTCGCG...CCGCGCGCGCGCGCGCGCGGTGAGCGG 263
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
264 CATTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGGGTCTCGGTGTC 313
101 AspLeuLeuLeuGluGlyLeuAlaGluAspSerIleAspSerIleLeuA 117
|||
314 GAA.....CCTGAAGGGATTGCTGAAGGTTCCATCGATAACACAGTAGT 357
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
|||||
358 TGTGCAAGTGAAGCAAGATTCTGAGATTGTGTTGGAAGGAGCAAGCTC 407
134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
408 GAGCTAAAGTAACACAAAACATTGCTTTTGAACGTGCGAAGCTTCTCT 457
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
458 TATGCAAGTCTGGGGGTCTAGGAGATTGTTGTTGTTTTCATTGCCAGTTGC 507
167 aLeuAlaAlaArgGlyHisArgValMetValMetValMetProArgTyrLeuA 184
|||||
508 TCTTGCTCTGCTGTCACCGTGTGATGTTGTAATGCCAGATATATTAA 557
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200

```

```
|||||.....|||||.....|||||.....|||||.....|||||
558 ATGCTACCTCCGATAAGAAATATGCAAAATGCAATTTACACAGAAAACAC 607
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217
|||||.....|||||.....|||||.....|||||.....|||||
608 ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCCATGA 657
217 uTyraGAspAsnValAspTrpValPheValAspHisProSerTyrrHisA 234
|||||.....|||||.....|||||.....|||||.....|||||
658 GTATAGAGATTTCAGTTGACCTGGGTGTTTGTGTGATCATCCCTCATATCACA 707
234 rgProGlySerLeuTyrrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||.....|||||.....|||||.....|||||.....|||||
708 GACCTGGAATTTATATATGGAGATAAGTTTGGTCTTTGGTGATAATCAG 757
251 PheArgTyrrThrLeuLeuCysTyrrAlaAlaCysGluAlaProLeuIleLe 267
|||||.....|||||.....|||||.....|||||.....|||||
758 TTCAGATACACACTCCTTTGCTATGCTGCGATGTGAGGCTCCTTTGGTCCCT 807
267 uGluLeuGlyGlyTyrrIleTyrrGlyGlnAsnCysMetPheValValAsnA 284
|||||.....|||||.....|||||.....|||||.....|||||
808 TGAATTTGGGAGGATATATTTATGACACAGAATTGCAATGTTTGTGTCATG 857
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrrArgPro 300
|||||.....|||||.....|||||.....|||||.....|||||
858 ATTGGCATGCCAGTCTAGTGGCAGCTCTTCTTGTGCAAAATATAGACCA 907
301 TyrrGlyValTyrrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||.....|||||.....|||||.....|||||.....|||||
908 TATGGTGTATAAAGACTCCCGGAGCATTCCTTGTATAACATAATTTAGC 957
317 aHisGlnGlyValGluProAlaSerThrTyrrProAspLeuGlyLeuProp 334
|||||.....|||||.....|||||.....|||||.....|||||
958 ACATCAGGGTGTAGAGCCTGCAACACATATCCTGACCTTGGTGGTGCAC 1007
334 roGluTrpTyrrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
1008 CTGAATGATATGAGCTCTGGAGTGGGTATTTCCCTGAATGGCGGAGGAGG 1057
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
1058 CATGCCCTTGACAGGGGTGAGCAGCTTAATTTTGAAGGTGCAGTTGT 1107
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrrSerTrpGluValT 384
1108 GACAGCAGATCGAATCGTGACTGTGCTGCTGAGTAAAGGTTATTTCATGGGAGTCA 1157
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1158 CAACTGCTGAAAGGTGGACAGGCCCTCAATGAGCTCTTAAGCTCCAGAAAG 1207
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
1208 AGTGATTAAAGGGAATGTAATGGAATTCACATTAAATGATTGGAACCC 1257
417 oThrThrAspLysCysLeuProHisHisTyrrSerValAspAspLeuSerG 434
1258 TGCCACAGACAAATGATATCCCTGTGCTTATTTCTGTTGATGACCTCTCTG 1307
434 lYlysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1308 GAAAGCCCAATGTAAGGTGCATTGCGAGAGGAGCTGGGTTCCTATATA 1357
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrrGlnLy 467
1358 AGGCCCTGATGTTCCCTCTGATTGGCTTTATTTGGAAGATTGGATTATCAGAA 1407
467 scLysIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
1408 AGGCCATTGATCTCATTTCACTTATCATCCAGATCTCATGCCGGGAGATG 1457
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||.....|||||.....|||||.....|||||.....|||||
```

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-12

seq_documentation_block:

; Sequence 12, Application US/08941445A

; Patent No. 6107060

; GENERAL INFORMATION:

; APPLICANT: Keeling, Peter

; APPLICANT: Guan, Hanping

; TITLE OF INVENTION: Starch Encapsulation

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,445A
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,855
FILING DATE: 30-SEP-1996
ATTORNEY/AGENT INFORMATION:

NAME: Winner, Ellen P
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089

SEQUENCE CHARACTERISTICS:
LENGTH: 1752 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant

MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORGANISM: Zea mays
FEATURE:

NAME/KEY: CDS
LOCATION: 1..1752
US-08-941-445A-12

alignment_scores:
Quality: 2713.50 Length: 585
Ratio: 4.916 Gaps: 3
Percent Similarity: 94.359 Percent Identity: 86.325

alignment_block:
US-09-674-824-2 x US-08-941-445A-12

Align seg 1/1 to: US-08-941-445A-12 from: 1 to: 1752

45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnG1 61
|||||
4 GTCGGGAGCTGAGCAGGAGGGGCCCGCCCGCCCGCTGCCACCCGC 53

61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78
54 GCTGCTGGCGCCCGCTCGTGGCCGGCTTCCTCGCG...CGCCGGCGC 100

78 laProAlaGlnSerProAlaProThrGlnProProLeuProAspAlaGly 94
|||||
101 AGCCACGGGTGAGCGGCATCCACCGCCCGCTGCCCGACCGCGC 150

95 ValGlyGluLeuAlaProAspLeuLeuLeuGluGlyLeuAlaGluAspSe 111
|||||
151 CTGGGGACCTCGGT.....CTCGAACCTGAAGGGATTCCTCAAGGTTTC 194

111 rleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMeta 128
|||||
195 CATCGATACACAGTAGTGTGGCAAGTGAGCAAGATTCTGAGATTGTGG 244

128 spAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPheVal 144
|||||
1145 CTGTTGATGACCTCTCTGAAAGGCCAAATGTAAAGGTGCAATTGCAGAAG 1194

245 TTGGAAGGAGCAAGCTCGAGCTAAAGTAACACAAAGCATTTGCTTTGTA 294
145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCy 161
|||||
295 ACCGGCGAAGCTTCTCCTTATGCAAGTCTGGGGGTCTAGGAGATGTTG 344
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValV 178
|||||
345 TGGTTTCATTCGCCAGTGTCTGCTGCTGCTGGTGGTGGTGGTGGTGGT 394
178 alMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAla 194
|||||
395 TAATGCCACAGATATTTAAATGGTACCTCCGATAAAGAAATATGCAATGCA 444
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisG1 211
|||||
445 TTTTACACAGAAAAACACATTCGGATTCCTTTGGCGGTGAACATGA 494
211 uValThrPhePheHisGluTyrArgAspAsnValAspTrpValPheValA 228
|||||
495 AGTTACCTTCTCCATGAGTATAGATTCAGTTGACTGGGTGTTGTTG 544
228 sPheHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGly 244
|||||
545 ATCATCCTCATATCACAGACCTGGAATTTATATGAGATTAAGTTGGT 594
245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystTyrAlaAlaCy 261
|||||
595 GCTTTTGGTGATAATCAGTTTCAGATACACACTCCTTTGCTATGCTGCATG 644
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnC 278
|||||
645 TGAGGCTCCTTTGATCCTTGAATGGGAGGATATATTTATGGACAGAA 694
278 ysMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294
|||||
695 GCATGTTTGTGTCATGATTCGATGCGCATGCCAGTCTAGTCCAGTCCCTCT 744
295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311
|||||
745 GCTGCAAAATATAGACCATATGTTGTTTATAAAGACTCCCGCAGCATTC 794
311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328
|||||
795 TGTAAATACATAATTTAGCACATCAGGGGTGTAGAGCCTGCAAGCACATATC 844
328 roAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGluTrpValPhe 344
|||||
845 CTGACCTTGGTGGTGGCCACCTGAATGCTATGGAGCTCTGGAGTGGGTATT 894
345 ProGluTrpAlaArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361
|||||
895 CCTGAATGGCGGAGGAGGATGCCCTTGACAAGGGTGAAGGCTGAATTA 944
361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnG 378
|||||
945 TTTGAAGGTGAGTGTGTACAGCAGATCGAATCGTACTGTCAAGTAAGG 994
378 lyTyrSerTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394
|||||
995 GTTATTCGTGGGAGGTCACAACCTGCTGAAGGTGGACAGGGCCTCAATGAG 1044
395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411
|||||
1045 CTCCTAAGCTCCAGAAAGAGTGTATTAACGGAATTTGTAATGGAATTGA 1094
411 pileAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428
|||||
1095 CATTAATGATTGGAACCTGCCACAGACAAATGTATCCCTGCTCATATT 1144
428 ervAlaAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
|||||
1145 CTGTTGATGACCTCTCTGAAAGGCCAAATGTAAAGGTGCAATTGCAGAAG 1194

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,951
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,602
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/263,921
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul N. Kokulis
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222957/1.02.15C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2990 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-572-951-1

alignment_scores:
    Quality: 2651.00      Length: 643
    Ratio: 4.587          Gaps: 7
    Percent Similarity: 89.891    Percent Identity: 79.471

alignment_block:
US-09-674-824-2 x US-08-572-951-1 ..
Align seg 1/1 to: US-08-572-951-1 from: 1 to: 2990

6 ValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAspProAl 22
613 CTCGCCGTCGGCGCGCGCTCCTCCTCGCGCGCGCTGCGCGGC 662
22 aThrAlaAla...ArgAlaSerAlaCysValValArgAlaLeuArgA 38
663 CGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700
38 rgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyProAla 54
701 CGGTGCTCGCGCGCGCGCGCGCTGCGCGGAGCTGAGCAGGGAGGCGCG 750
55 AlaArgProAlaGlnGlnGlnGlnLeuAlaProProLeuValProGlyph 71
751 CGCGCGCGCGCTGCCACCGCGCTGCTGCGCGCGCGCGCGCGCTG 800
71 eleuAlaProProProAlaProAlaGlnSerProAlaProThrGlnP 88
801 CCTCGCG...CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 847
88 toProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeuLeu 104
848 CGCGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 891
105 GluGlyTlleAlaGluAspSerIleAspSerIleValAlaAlaSerG1 121
892 GAAGGGAATGCTGAAGGTTCATCATCAACACAGTAGTTGTGGCAAGTGA 941
121 uGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAlaLysVal 138
942 GCAAGATTCTGAGATTGTGTTGGAAAGGAGCAAGCTCGAGCTAAGTAA 991

;
; seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-572-951-1
;
; seq_documentation_block:
; Sequence 1, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
; TITLE OF INVENTION: MODIFICATION OF STARCH
; TITLE OF INVENTION: SYNTHESIS IN PLANTS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: Pillsbury Madison & Sutro LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```
138 hrArgSerIleValPheValThrGlyGluAlaPro.TyrAlaLysSe 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
992 CACAAAGCATCTCTTTGTAACGGCGAGCTTCCTTAATCGAAAGTC 1041
154 rGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaA 171
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1042 TGGGGGCTCTAGAGAGATGTTGTGGTTCATTGCCAGTGTCTCTGTGCTC 1091
171 rGlyHisArgValMetValValMetProArgTyrLeuAsnGlySer 187
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1092 GCGGTCACCGTGTGATGTTGTAATGCCAGACATTTAAATGGTACCTCC 1141
188 AspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIle 204
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1142 GATAAGAAATTATGCAATGCAATTTTACTCAGAAAACACATTCGGATTCC 1191
204 ocysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgAsp 221
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1192 ATTCCTTTGGCGGTGAACATGAAGTTACCTTCTTCATGAGTATAGAGATT 1241
221 snValAspTrpValPheValAspHisProSerTyrHisArgProGlySer 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1242 CAGTTGACTGGGTGTTGTTGATCATCCCTCATATCAGACACCTGGAAAT 1291
238 LeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrTh 254
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1292 TTATATGAGATAAGTTTGGTGTCTTTTGGTGAATAATCAGTTTCAGATAC 1341
254 rLeuLeuCysTyrAlaAlaCysGluAlaProIleLeuLeuGlyGly 271
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1342 ACTCTTTTGTCTATGTCGATGTGAGGCTCCTTTGGTCTCTGAATGGGAG 1391
271 lyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHisAla 287
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1392 GATATATTTATGGACAGAAATGTCATGTTGGTTGTCATGATGTCATGCC 1441
288 SerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTy 304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1442 AGTCTAGAGCCAGCTCTCTTCTGTCGCAAAATATAGACCATATGCTGTTA 1491
304 rArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGly 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1492 TAAAGACTCCCGCAGCATCTCTGTAATACATAATTTAGCACATCAGGGTG 1541
321 alGluProAlaSerThrTyrProAspLeuGlyLeuProProGluTrpTyr 337
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1542 TAGAGCCTGCAAGCACATATCCTGACCTTGGGTTGCCACCTGAATGCTAT 1591
338 GlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAs 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1592 GGACCTCTGGAGTGGGTATTCCCTGGAATGGCGAGAGGATGCCCTTGA 1641
354 pLysGlyGluAlaValAsnPheLeuLysGlyValAlaValThrAlaAspA 371
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1642 CAAGGCTGAGGACGTTAAATTTTGAAGGTGCAGTTGTCACAGCAGATC 1691
371 rGlyValThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGlu 387
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1692 GAATCGTGACTGTGTAAGGGTATTTCATGGGAGGTCAACAACCTGCTGAA 1741
388 GlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAs 404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1742 GGTGGACAGGCCCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATATA 1791
404 nGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspL 421
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1792 CGGAATTTGTAATGGAATTCACATTAATGATTGGAACCTGCCACAGACA 1841
421 yscysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAlaLys 437
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1842 AATGTATCCCTGTCTATTTCTGTTGATGACCTCTCTTGAAGAGGCTAAA 1891
438 cysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspVa 454
```

```
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1892 TGTAAAGGTGATTGCAGAGGAGCTGGGTTTACCTATAAGGCCTGTGATG 1941
454 lProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspL 471
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1942 TCCCTGATTGGCTTTATTGGAAGATTGGATTATCAGAAAGGCATTGATC 1991
471 euIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPheVal 487
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1992 TCATTCAACTTATCATACCAGATCTCATCGGGAAGAATGTTCAATTTGTC 2041
488 MetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGl 504
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2042 ATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGA 2091
504 uSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValProV 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2092 GTCGATCTTCAAGCATAAATTTCTGTGGATGGGTTGGATTAGTGTTCAG 2141
521 alSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArg 537
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2142 TTTCCCAACCGAATAACTCGGCTGGCGATATATTGTTAATGCCATCCAGA 2191
538 PheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrVa 554
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2192 TTCGAACCTTGTGCTCTCAATTCAGCTATATGCTATGCAATGTCACAGT 2241
554 lProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheA 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2242 TCCCTGTTGTCATGCAACTGGGGCCCTTAGAGATACCGTGGAGAACTCA 2291
571 snProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerPro 587
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2292 ACCCTTTCGGTGAGAAATGGAGCAGGCTACAGGGTGGGCATTCGCACCC 2341
588 LeuThrValAspLysMet.LeuTrpAlaLeuArgThrAlaMetSerThrP 604
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2342 CTAACACACAGAAACATGTTTGTGGACATTCGCAACTGCAATATCTACAT 2391
604 heArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThr 620
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2392 ACAGGGAACACAAAGTAATAATGGGAGGGCTAATGAAGCCAGGCATGTCA 2441
621 LysAspHisThrTrpAspHisAlaPro.SerSerThrSerArgSerSerS 637
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2442 AAGAGTTCACGTGGGACCATCGCGTGAACAATACGAAACAAATCTCCA 2491
637 erGlyProSerTrpTrpAsnPro 644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2492 GTGGGCTTCATCGGATCGACCC 2514
```

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-20

```
seq_documentation_block:
; Sequence 20, Application
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Wimmer, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1620 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL: NO
: NAME/KEY: CDS
: LOCATION: 1..1620
: US-08-941-445A-20

alignment_scores:
  Quality: 2546.50      Length: 538
  Ratio: 4.935          Gaps: 3
  Percent Similarity: 95.911 Percent Identity: 87.918

alignment_block:
US-09-674-824-2 x US-08-941-445A-20 ..

Align seg 1/1 to: US-08-941-445A-20 from: 1 to: 1620

95 ValGlyGluLeuAla...ProAspLeuLeu.....GluGlyIleAl 108
   |||||
   4 GTCCGGGAGCTGAGCAGGAGGAGCCTCGCTCGAACCTGAAGGATGC 53

108 aGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAspSerG 125
   |||||
   54 TGAAGTTCCATCGATAACACAGTAGTTGTGGCAAGTCAGCAAGATTCTG 103

125 luIleMetAspAlaasnGluGlnProGlnAlaLysValThrArgSerIle 141
   |||||
   104 AGATTGTTGGTGGAAAGGAGCAAGCTCGAGCTAAAGTAACACAAAGCAT 153

142 ValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGl 158
   |||||
   154 GTCTTGTAAACCGGGAAGCTTCTCTTATGCAAAAGTCTGGGGGTCTAGG 203

158 yAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgV 175
   |||||
   204 AGATGTTGTGTTGCTCATTGCCAGTTGCTCTTCTGCTGCTGCTACCCGTG 253

175 aIleMetValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyr 191
   |||||
   254 TGATGTTGTTAATGCCAGATATTAAATGGTATGATCCGATCCGATTAAT 303

192 AlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGl 208
   |||||
   304 GCNAATGCATTTTACACAGAAACACACATTTCGATTCCTGCTTTGGCGG 353

208 ySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTpv 225
   |||||
   354 TGAACATGAAGTAGTACCTTCTCCATGAGTATAGAGATTACAGTTGACTGGG 403

225 aIlePheValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAsp 241
   |||||
   404 TGTGTTGTCATCATCCCTCATATACAGACCTGGAAATTTATATGAGAT 453

242 AsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystY 258
   .....
454 AGTTTGGTCTTTTGGTGATATAGTTTCCAGATACACACTCCTTTGCTA 503

258 rAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrG 275
   .....
504 TGCTGCATGTGAGGCTCCTTTTGATCTTGAATTTGGGAGGATATATTATG 553

275 lYGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuValPro 291
   .....
554 GACAGAATTGCATGTTTGTGTCAATGATTGGCATGCCACTCTAGTCCCA 603

292 ValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerAr 308
   .....
604 GTCCTTCTTCTGCTCAAAATATAGACCATATGGTGTATATAAGACTCCCG 653

308 gSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlas 325
   .....
654 CAGCATTCTTTGTAATACATAATTTAGCACATCAGGGGTAGAGCCCTGCAA 703

325 erThrTyrProAspLeuGlyLeuProGluTyrProGluTyrGlyAlaLeuGlu 341
   .....
704 GCACATATCTGACCTTGGGTGCCACCTGAATGGTATGAGCTCTGGAG 753

342 TrpValPheProGluTyrAlaArgArgHisAlaLeuAspLysGlyGluAl 358
   .....
754 TGGGTATTCCCTGAATGGCGAGGAGGCATGCCCTTGACAAGGGTGAGGC 803

358 aValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrv 375
   .....
804 AGTTAAATTTTGAAGAGGTGCAGTTGTGACAGCAGATCGAATCGTACTG 853

375 alSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGlnGly 391
   .....
854 TCAGTAAGGGTTATTCTGGGAGGCTCACAACTGCTGAAGGTGGACAGGC 903

392 LeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValas 408
   .....
904 CTCAATGAGCTCTTAAAGCTCCAGAAAGAGTGATTAACCGAATTTGTA 953

408 nGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuProH 425
   .....
954 TGGAAATGACATTAATGATGGAACCCCTGCCACAGACAAATGTATCCCT 1003

425 isHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGlu 441
   .....
1004 GTCATTATTCTGTTGATGACCTCTCTGAAAAGGCCAAATGTAAAGTGCA 1053

442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleGl 458
   .....
1054 TTGCAGAAAGGAGCTGGGTTTACCTATAGGCGCTGATGTTCTCTGATTGG 1103

458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMeta 475
   .....
1104 CTTTATGGAAGGTGGATTATCAGAAAGGCATTCATCTCATTCNACTTA 1153

475 laIleProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491
   .....
1154 TCATACCAGATCTCATGGGGAAGATGTTCAATTTGTCATGCTTGGATCT 1203

492 GlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLy 508
   .....
1204 GGTGACCACAGCTTGAAGATTTGAGATGAGATCTACAGAGTCGATCTTCAA 1253

508 sAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArgI 525
   .....
1254 GCATAAATTTCTGTCGATGGTGGATTAGTGTCCAGTTTCCCACCCAA 1303

525 leThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCys 541
   .....
1304 TAACCTGCGGCTGCGATATATTGTTAATGCCATCCAGATTCGAACCTTGT 1353
```



```

542 GlyLeuAsnGlnLeuThrAlaMetGlnTyrGlyThrValProValValHi 558
1354 GGTCTCAATCAGCTATATGCTATGAGTATGCGACAGTTCCTGTGTGCTCA 1403
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575
1404 TGCAACTGGGGCCCTTAGAGATACCGTGGAGAACTTCAACCTTTCGGGTG 1453
575 laLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValAsp 591
1454 AGAATGGAGACAGCGGTACAGGGTGGGCATTCGCCACCCCTAACACAGAA 1503
592 LysMet.LeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisL 608
1504 AACATGTTTGTGGACATTCGCAACTGCAATATCTACATACAGGGAACACA 1553
608 ysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 624
1554 AGCTCTCTGGGAAGGGCTAATGAAGCGAGGCATGTCAAAAGACTTTCACG 1603
625 TrpAspHisAla 628
1604 TGGGACCATGCC 1615

```

seq_name: /cgn2_5/ptodata/1/ina/6A_COMB.seq:us-08-836-567-9

```

seq_documentation_block:
; Sequence 9, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Koosmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

```

; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Dshire
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in Lambda ZAPII
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 68..1990
; US-08-836-567-9

```

alignment_scores:
Quality: 2001.00 Length: 545
Ratio: 4.222 Gaps: 5
Percent Similarity: 86.972 Percent Identity: 66.789

alignment_block:
US-09-674-824-2 x US-08-836-567-9 ..

Align seg 1/1 to: US-08-836-567-9 from: 1 to: 2360

```

102 LeuLeuLeuGluGlyIleAlaGluAspSer.....IleAspSerIleI 116
116 eValAlaAspSerGluGlnAsp.....SerGluIleMetAspAlaAsnG 131
335 TTGATTCACATCTCTGTGAGGAGATGCAACAATGGTAGAATCTCATGA 384
385 TATTGTAGCCAATGATAGAGATGACTTGAGTGGAGGATGAGGAGATGG 434
131 luGlnProGlnAlaLysValThrArgSerIleValPheValThrGlyGlu 147
435 AGGAACCCCAATCAATTAACATTCATATCATTTTGTACTGCTGAA 484
148 AlaAlaProTyrAlaLysSerGlyGlyLeuGlyValCysGlySerLe 164
485 GCAGCTCCATATCTAAGACTGGTGGATTAGGAGATCTTTGTGGTCTTT 534
164 uProIleAlaLeuAlaAlaArgGlyHisArgValMetValMetProA 181
535 GCCAATGGCAGTACGTCTCGGGGTGATCGTGAATGGTCTTTCACCTA 584
181 rGlyrLeuAsnGlySer...SerAspLysAsnTyrAlaLysAlaLeuTyr 196
585 GGTATTGATGAGGTCTTCAGATGAAAGTACGCCAATGCTGTGAC 634
197 ThrAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValTh 213
635 CTTGATGTGGGGCCACTGTCCATTGCTTTGGTGTGATGCACAGGAGTAGC 684
213 rPhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHisP 230
685 CTTCTACCATGAATACAGGCGAGGTGTGATTGGGTATTGTGGACCACT 734
230 roSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPhe 246
735 CTTCTTACTGCAGACCTGGAAACCCCATATGGTATATTATGGTGCATTT 784
247 GlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAl 263
785 GGTGATAATCAGTTTCGCTTCACTTTGCTTCTCAGCAGCATGTGAAGC 834
263 aProLeuIleLeuGluGlyTyrIleTyrGlyGlnAsnCysMetP 280
835 GCCATTGGTCTTCCACTGGGAGGGTTCACTTATGGAGAGAAAGTCTTGT 884
280 heValValAsnAspTrpHisAlaSerLeuValProValLeuAlaAla 296
885 TTCTCGCTAATGATGGCATGCTGCCCTGTTCCITTTACTTTTAGCGGCC 934
297 LysTyrArgProTyrGlyValTyrArgSerArgSerThrLeuValI 313
935 AAGTATCGTCTTATGGTGTTTACAAGGATGCTCGTAGTATTGTCGAAT 984

```

```

612 uGlyLeuMethlysArgGlyMetThrIysAspHisThrTrpAspHisala. 620
    ||| ::: :||| ||| :||| :||| :||| :||| :||| :|||
1885 GGGATTGATGAGGAGGAGTGGGAAGGACTATTCCTGGGAAATGCAG 1934

629 ProSerSerThrSerArgSerSerSerGlyPro 639
    ||| ::: :||| ||| :||| :||| :||| :||| :||| :|||
1935 CCATTCAATATGAACAAGTTTTCACCTGGGCGCT 1967

seq_name: /cgn2_6/ptodata/1/1na/6a_COMB.seq:US-08-836-567-3

seq_documentation_block:
; Sequence 3, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrovo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1758 base pairs
; TYPE: nucleotide
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Berollina
; TISSUE TYPE: tuber tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in pBluescriptSKII+
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1377
; OTHER INFORMATION: /function= "Polymerization of
; OTHER INFORMATION: starch"
; OTHER INFORMATION: /product= "Starch synthase"
; US-08-836-567-3

```

```

alignment_scores:
  Quality: 1731.50      Length: 528
  Ratio: 4.142          Gaps: 7
  Percent Similarity: 79.167  Percent Identity: 61.364

alignment_block:
  US-09-674-824-2 x US-08-836-567-3 ..

Align seg 1/1 to: US-08-836-567-3 from: 1 to: 1758

203 lleProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrAr 219
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
40 GTCCATTGCTTGTGTGATGCACAGAGTAGCCTTCTACCATGAATACAG 89

219 gAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 236
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
90 GGCAGGTGTTGATGGGTATTTGGACCACCTCTCTTACCGCAGACCTG 139

236 lySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArg 252
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
140 GAAGCCCATATGTTGATATTTATGTCATTTGGTGAATACAGTTTCGC 189

253 TyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuLeuLeuLe 269
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
190 TTCATTGCTTCTCAGCGCAGCATGTAAGCGCCATTTGGTCTTCCACT 239

269 uGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpH 286
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
240 GGGAGGTTTCACTATGAGAGAGAGTCTGTTTCTCGCTAATGATGCA 289

286 lsAlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrGly 302
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
290 ACGTGGCTTGGTTCCTTTATTTTAGCGGCCAAGTATCGTCCTTATGGT 339

303 ValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
340 GTTTACAGAGTCTCTAGTATTGTCGCAATACACAACTATGACATCA 389

319 nGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProGluT 336
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
390 GGGAGTGGAGCTCGAGTAACCTACAATAATTTGGGTTTGCCTCCACA 439

336 rpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAl 352
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
440 GGTATGAGCAGTTGAATGGATATTTCCACATGGCAAGGCGCATGCG 489

353 LeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValThrAl 369
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
490 CTTGACACTGGTGAACAGTGAACGTTTGAAGGGGCAATAGCAGTTGC 539

369 aAspArgIleValPheValSerGlnGlyTyrSerTrpGluValThrTh 386
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
540 TGATCGGATACTGACACTGACCGAGGATCATCGGAAATACAACTC 589

386 laGluGlyGlyGlnGlyLeuAsnGlnLeuLeuSerSerArgLysSerVal 402
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
590 CTGAAGGGGATATGGGCTACATGAGCTGTTGAGCAGTAGACAGCTGTT 639

403 LeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 419
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
640 CTTAATGGAATTAATGAATAGATGTTAATGATTGGAACCGCTGCAC 689

419 rAspLysCysLeuProHisTyrSerValAspAspLeuSer.....G 434
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
690 AGATGACATATCGCTTCGATTAATGCAATGCAATGACCTCTCCCCCTG 739

434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
740 GAAAGGTTTCAGTCAGACAGTATGTCGCAAAAGGAACCTGGGCTTCCA 789

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGln 467
   ||| ||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

790 CGACCCGATTGTCCTCACTGATTGGATTATTTGGAAGGCTGGACTACAGAA 839
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
840 AGGTGTTGACATAATCCTGTCAGCAATTCAGAACTTATGCAAGATGATG 889
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
890 TCCAAGTTGTAATGCTTGGATCTGCTGAGAAACAATATGAAGACTGGATG 939
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
940 AGACATACAGAAATCTTTTAAAGACAAATTTCTGCTGGGTGGATT 989
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeu 534
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
990 TAATGTTCCAGTTTCTCATAGTAACAGCAGGATGCACATATATTGA 1039
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1040 TGCCTCAAGATTCAACCGGTGCTTAAACCAATTTGATGCAATGAGA 1089
551 TyrGlyThrValProValHisGlyThrGlyLeuArgAspThrVa 567
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1090 TATGCCACCATACCTATTGTTTCATAGCAGCGGGGSCCTAAGAGACACAGT 1139
567 lGluThrPheAsnProPheGlyAlaLysGlyGlu...GluGlyThrGlyT 583
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1140 GAAGGATTTTAATCCATATGCTCAAGAAAGGAAAGGTGAAGGTACCGGT 1189
583 rpAlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThr 599
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1190 GGACATTTTCTCTCAACGAGTGAAGTTGTTGTATACATCACTGAAGCTG 1239
600 AlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLy 616
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1240 GCGATCAGGACTTATACAGAACATAAGTCATCTTGGGAGGAGTTGATGAA 1289
616 sArgGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThr 632
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1290 GAGAGGTATGGAGGAGCTATTCCTGGGAAATGACAGCATTCATCAATATG 1339
633 SerArgSerSerGlyPro..... 639
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1340 AGCAAGTTTTCACCTGGGCTTTTATAGATCTCCATATGTCAGATGATTT 1389
640 .....SerTrpThrAsnProThrSerCysArgArgGlyL 651
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1390 ATCAAGAAAGATTGCAAAACGGGATACATCATTAACATATACGAGAGCTT 1439
651 euGly.....ArgSerLysCysGluSerPro 659
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1440 TTGGTGTATTAGTACTGTCTGTCATTTGGCGCGGAATGTTGT..... 1480
660 SerAlaLeuLysThrSerSerSerPheArgGlyProGluGlyTyrPr 676
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1481 .....GTTCTTCTGATTTCAGAGAGATCAAGTAGTCTCC 1515
676 oCysThrLeuArgCysProAlaThrVal..... 685
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1516 AAAGACATGTAGCTGCGCTGCTGTGTGATGAAGTAAACTACAAAGGCA 1565
686 .....GluSerGlnCysAlaCysLeuLeuTrp 694
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1566 ATTAGAAACCCCAACAACTGCTCTCTTTGG 1597

```

```

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-10
seq_documentation_block:
; Sequence 10, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter

```

APPLICANT: Guan, Hanning
 TITLE OF INVENTION: Starch Encapsulation
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 STREET: 5370 Manhattan Circle
 CITY: Boulder
 STATE: CO
 COUNTRY: US
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/941,445A
 FILING DATE: 30-SEP-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,855
 FILING DATE: 30-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Winner, Ellen P
 REGISTRATION NUMBER: 28,547
 REFERENCE/DOCKET NUMBER: 89-97
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2097 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: not relevant
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2097
 US-08-941-445A-10

alignment_scores:
 Quality: 1246.00 Length: 714
 Ratio: 2.800 Gaps: 24
 Percent Similarity: 62.325 Percent Identity: 42.857
 alignment_block:
 US-09-674-824-2 x US-08-941-445A-10 ..
 Align seg 1/1 to: US-08-941-445A-10 from: 1 to: 2097
 7 GlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAspProAlaTh 23
 |||||
 65 GCGCAGCGCGG...GCAGTGGGTGCTCTCCGCTCGCTACGCGCTAC 111
 23 rAlaAlaArgAlaSerAlaCysValValArg.....AlaArgLeuArgA 38
 |||||
 112 ACGGCGCGGAGCTCGGTGATTGGGCGGCGGCGGCGGCGCTC...A 158
 38 rGLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyProAla 54
 |||||
 159 GGATGAGCGCGGTGGTACGCGCGCGCAGCGCGCGCGCGGCGGAA 208
 55 AlaArgProAlaGlnGlnGlnLeuAlaProPro...LeuValPro.... 69
 |||||
 209 GCGAGG.....AGGCAGCGAGAGCTCTCTCTCTCTCTCTCTCTCT 246
 70GlyPhe...LeuAlaProPr 75
 |||||

247 GCGCTGTTTCAGGCGACGACGCGCAAGGCTGTGGATTCTGCTTCACTCC 296
 75 oProPro.....
 |||||
 297 CAATCCTTTGACATCTCTCGGAAGCAAAATCAGCGCTGCAATGCAAA 346
 78AlaProAlaGlnSer 82
 |||||
 347 ACGAACGAGTGGGGCAGCAGCGCGAGCAGCGCGCGCGGTGTCGGA 396
 83 ProAlaProThrGlnProProLeuPro.....AspAl 93
 |||||
 397 CCCAAGCTGATCATCTCAGCTCTGTGTCCACCAAGAGAGAAATCGATGC 446
 93 aGly.....ValGlyGluLeuAlaProAspLeuLeuL 104
 |||||
 447 CAGTCCGCTGAGCCAGCGCGCGAGCTGATGCTAGACCGGTGCAAA 496
 104 euGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaAlaSer 120
 |||||
 497 GCATAGGCATCGCTGAA...CCGGTGGATGCTAAGGCTGATGACGCTCCG 543
 121 GluGlnAspSer.....GluIleMetAspAlaAsnGl 131
 |||||
 544 GCTACAGATCGCGCGCGAGTGTCTTATGACAGGAGGAGATAATGAACC 593
 131 uGlnProGlnAla...LysValThrArgSerIleValPheValThrGlyG 147
 |||||
 594 TGGCCCTTGGCTGGGCTAATGTGATGAAGCTGCTGCTGGTGGCTTCTG 643
 147 luAlaAlaProTyrAlaLysSerGlyLeuGlyAspValCysGlySer 163
 |||||
 644 AATGTGCTCTTCTGCAAGCAGGTGGCTTGGAGATGCTGGTGGTGGCT 693
 164 LeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMetPr 180
 |||||
 694 TTGCTAAGGCTCTGGCAGGAGGAGACCGCTGTATGCTGCTGATACCC 743
 180 oArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrT 197
 |||||
 744 AAGATAT.....GGAGAGTATGCGGAGCGCGGATT 775
 197 hrAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValThr 213
 |||||
 776 TAGGTGAAGGAGAGCTTACAGGTAGCTGGCAGGAGATTTCAGAGTTACT 825
 214 PhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHisPr 230
 |||||
 826 TATTTTCACTCTTACATTTGATGGATTTGATTTGTTATTCGTAGAAGCCC 875
 230 oSerTyr...HisArgProGlySerLeuTyrGlyAspAsnPheGlyAlap 246
 |||||
 876 TCCCTTCCGCGCGCGCACATAATATTTATGGG.....GGAGAAA 916
 246 heGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGlu 262
 |||||
 917 GATTGGATATTTGAGCGCATGATTTTGTCTGCAAGCGCGCTCTTGAG 966
 263 AlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln...AsnC 278
 |||||
 967 GTTCCATGGTATGCTCCACTGTGGCGTACTGTCTATGTTGATGGCAACT 1016
 278 smetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeuA 295
 |||||
 1017 AGTTTTCATTTGCTAATGTCATACCGCACTCTCTGCTGTCTATCTAA 1066
 295 laAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLeu 311
 |||||
 1067 AGGCCTATACCGGACCAATGTTTGTGATGACAGTATGCTGCTCTGCTT 1116
 312 ValIleHisAsnLeuAlaHisGlnGlnValGluProAlaSerThrTyrPr 328
 |||||
 1117 GTGATACACAACATTGCTCATCAGGCTGTGGCCCTGTAGACGACTTCGT 1166

328 oAspLeuGlyLeuProGluTrpTyrGlyAlaLeuGluTrpValPheP 345
: : : : :
1167 CAATTTTGACTTGCTT 1182
345 roGluTrpAlaArgArgHisAlaLeuAspLys 355
1183GAACTACTGACCTGACCACTTCAAACTGTATGACAAC 1218
356GlyGluAlaValAsnPheLeuLysGlyAlaValValThrAlaAs 370
1219 ATTGTGGGATCAGACACGCTTTTCTCGGGGCTGAACACGGCAGA 1268
370 pArgLeuValThrValSerGlnGlyTyrSerTrpGluValThrAlaG 387
1269 CCGGTGTGACCGTTAGCATGTGATGCTGAGCTGAAGACTTCGG 1318
387 luGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysValLeu 403
1319 AAGCGGGTGGGCTCCACACATCATAAACAGACGCTGGAAGCTG 1368
404 AsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrAs 420
1369 CAGGCATCGTGAACGGCATCGACATGAGCGAGTGGAAACCCGCTGTGA 1418
420 pLysCysLeuProHisHisTyrSerValAspAspLeu. 432
1419 CGTGCACCTCCACCTCCGACGACTACACCACTACACGTTCCGAGCGCTG 1468
433 ..SerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeu 448
1469 ACACCGGCAACGGCAGTGCAGGCGCCCTGCGAGCGGCTGGGCTG 1518
449 ProValArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTy 465
1519 CAGGTCCGCGACGAGCTGCCACTCATCGGTTCATCGGCGGCTGGACCA 1568
465 rGlnLysGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArg 482
1569 CCAGAGGGGTGGACATCATCGCGGACGCGATCCACTGGATCGCGGGC 1618
482 luAspValGlnPheValMetLeuGlySerGlyAspProIlePheGluGly 498
1619 AGGACGTGCAGCTCGTGCATGCGGCACCGCGGGCGGCGGCTGGAGGAC 1668
499 TrpMetArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpVa 515
1669 ATGCTGGCGGGTTCGAGTCCGAGCAGCAGCAGCAAGGTGCGCGCTGGT 1718
515 lGlyPheSerValProValSerHisArgIleThrAlaGlyCysAspIleL 532
1719 GGGGTTCTCGTCCCTGCGGACCGCATCAGCGGGGCGGCGGATCC 1768
532 euLeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAla 548
1769 TGCTGATGCGCTCGCGGTTCGAGCGGTGCGGGGTGAACACGACTACGCC 1818
549 MetGlnTyrGlyThrValProValValHisGlyThrGlyCysLeuArgAs 565
1819 ATGGGTACGGGACCGTGGCGGTGGTGCACGCGGTGGGGGGTCCGGGA 1868
565 pThrValGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrG 582
1869 CACGGTGGCGGCTTCGACCCGTCAACGACCGGGCTCG 1909
582 lYTrpAlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArg 598
1910 GGTGACGTTGACCGCGCGGAGGCGAACCAGGATGATCGACGCTCTCG 1959
599 ThrAlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMe 615
1960 CACTGCGCTCACCACGCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2009

615 tLysArgGlyMetThrLysAspHisThrTrpAspHisAla.Pro..... 629
: : : : :
2010 GCGCGCGCATGCGCGAGGACCTCAGCTGGACCGCCGCTGCTGT 2059
630SerSerThrArgSerSerSerGly 638
2060 ATGAGGACGCTCTGCTCAAGCGGAGTACCAGTGTGT 2095
seq_name: /cgn_6/ptodata/1/lna/6A_COMB.seq:US-08-836-567-7
seq_documentation_block:
: Sequence 7, Application US/08836567
: Patent No. 6130367
: GENERAL INFORMATION:
: APPLICANT: Kossmann, Jens
: APPLICANT: Springer, Franziska
: APPLICANT: Abel, Gernot
: TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
: TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
: TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FISH & NEAVE
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/836,567
: FILING DATE: 24-JUL-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP95/04415
: FILING DATE: 09-NOV-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE P 44 41 408.0
: FILING DATE: 10-NOV-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Haley Jr., James F.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: Agrevo-4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-596-9000
: TELEFAX: 212-596-9090
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2793 base pairs
: TYPE: nucleotide
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Solanum tuberosum
: STRAIN: cv Desire
: TISSUE TYPE: leaf tissue
: IMMEDIATE SOURCE:
: LIBRARY: cDNA-library in Lambda ZAPII
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 242..2542
: US-08-836-567-7

alignment_scores:

Quality: 1215.00 Length: 578
Ratio: 3.030 Gaps: 12
Percent Similarity: 69.377 Percent Identity: 45.675

339 laLeuGluTrpValPheProGluTrpAlaAcrArgHISalalaLeuAspLys 355
: : : : : |||
1647 CTTTCAAGTTGTATGACCCAGTA.....GGA 1672

356 GlyCluAlaValAsnPheLeuLysGlyAlaValThrAlaAspArgIle 372
: : : : : |||
1673 GGTGAGCATTTCAACATTTTTCGGCTGGTCTAAAGACGACGATCGTGT 1722

372 eValThrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyG 389
: : : : : |||
1723 AGTTACAGTTAGTCATGGATATTATCATGGCAACTAAAGACTTCCCAAGGTG 1772

389 lyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGly 405
: : : : : |||
1773 GTTGGGATTTGCATCATAGATAAATTAATGAGAACGATTTGGAAATACACGGT 1822

406 IleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysC 422
: : : : : |||
1823 ATTGTGNAATGGGATTGATACAAAGAGTGGAAACCCCTGAGTTGGACGTTCA 1872

422 sLeuProHis.....HisTyrSerValAspAspLeu...SerG 434
: : : : : |||
1873 CTTACAGTCAGATGGTTACATGAACACTCTCTTGGACAGCTACAGACTG 1922

434 LysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
: : : : : |||
1923 GCAAGCCTCAATGTAAAGCTGCATTCGACAGAGGAACCTGGTTACCACTT 1972

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnL 467
: : : : : |||
1973 CGTCATGATGCCACATGATCGGTCTTTCATTTGGGAGGCTTGACCCACAAA 2022

467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
: : : : : |||
2023 GGGTGTGATCTGATCTGCTGAGGCCACTGCTTGGATGATGGTTCAGATG 2072

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrPmet 500
: : : : : |||
2073 TACAACCTGGTCATGTTGGGACGGGAGCGGTGCACCTTGAACAGATGCTA 2122

501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
: : : : : |||
2123 AGGCAATTTGAGTCTCAACACAAATGATAAATTAGAGATGGGTGGTGT 2172

517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
: : : : : |||
2173 CTCTGTGAAGACTTCTCATCTGATACTGCTGGTGGCAGACATCTGCTCA 2222

534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
: : : : : |||
2223 TGCCCTCTAGATTTGAGCCTTTCGCGACTGAACCAACAGCTTTATGCAANTGAA 2272

551 TyrGlyThrValProValValHisGlyThrClyGlyLeuArgAspThrVa 567
: : : : : |||
2273 TATGGGACTATTCCTGTTGTTTCATGCGACTAGGAGGACTCAGAGACTGT 2322

567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
: : : : : |||
2323 GCAGCCCTTTGATCCTTT.....AATGACTCAGGACTGGGGTGGGA 2363

584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
: : : : : |||
2364 CCTTCAGTAGGGCTCAAGCTAGCCAGCTGATCCACGCGATTAGGAATATGC 2413

601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
: : : : : |||
2414 TTACTGACTATTCGTGTAGTACAAAAGACTTTGGAGGGGATTCAGACACG 2463

617 gGlyMetThrLysAspHisThrTrpAspHisAla 628
: : : : : |||
2464 TTGCTATGACACAGACTTAAAGTTGGTGGATAATGCT 2497


```
825 TTGGCGGTGGTCTAAACACACAGCATCGTGTAGTTACAGTTAGTCATGGA 874
379 TyrSerTrpGluValThrThrAlaGluGlyGlyClnGlyLeuasnGluLe 395
875 TATTTCATGGGAACATAAAGACTTCCCAAGGTGGTGGGATGTCATCAGAT 924
395 uLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspI 412
925 AATTAATGAGAACGATTGGAAATGAGGGGTATTGTGAATGGGATGATA 974
412 leAsnAspTrpAsnProThrThrAspLysCysLeuProHis..... 425
975 CAAAAGAGTGAACCCCTGAGTTGACGCTTCACTTACAGTCAGATGGTTAC 1024
426 ...HisTyrSerValAspAspLeu...SerGlyLysAlaLysCysLysAl 440
1025 ATGAACACTACTCTTGGACACCGCTACAGCTGGCAAGCCTCAATGTAAGC 1074
440 aGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI 457
1075 TGCATTGCAGGAAGCAACTGGTTTACCAGTTCGTGATGATGCCACTGA 1124
457 leGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLys 473
1125 TCGGTTTCATTGGGAGGCTTGCACCAACAAAAGGTTGTGATCTGATTGCT 1174
474 MetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuG 490
1175 GAGGCCACTGCTTGGATGATGGTCAGATGTACAACCTGGTCATGTTGGG 1224
490 ySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSert 507
1225 GACGGGGAGGCGTGACCTTGAACAGATGCTAAGGCAATTTGAGTGCTAAC 1274
507 yLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHis 523
1275 ACAATGATAAAATTAGAGATGGGTGGTTGCTTCTGTGAAGACTTCTCAT 1324
524 ArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPr 540
1325 CGTATAACTGCTGGTGGACAGACATTCGTGCTCATGCCCTTCTAGATTGAGGC 1374
540 oCysGlyLeuAsnGlnLeuTyAlaMetGlnTyGlyThrValProValV 557
1375 CTTCGCACTGAACCAAGCTTATGCAATGAAATATGGGACTATTCTGTTG 1424
557 alHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPhe 573
1425 TTTCATGCAGTAGGAGACTCAGAGATATGTCGACCCCTTTGATCCTTTT 1474
574 GlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThrVa 590
1475 .....AATGAGTCAGGACTGGGTGGGCTTCACTTCACTAGGCTGAAGC 1515
590 lAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607
1516 TAGCCAGCTGATCCAGCATTAAGGAATTCCTTACTGACTTATCTGAGT 1565
607 iSLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
1566 ACAAAAAGAGTTGGGAGGGGATTGACACACGTTGTATGACACAGACTTA 1615
624 ThrTrpAspHisAla 628
1616 AGTTGGGATAATGCT 1630
```

```
seq_name: /cgn2_6/ptodata/1/1na/6B_COMB.seq:US-09-196-390-5
seq_documentation_block:
; Sequence 5, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina

; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOPHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum L.
; STRAIN: cv. Florida
; TISSUE TYPE: ca. 21 d Caryopses
; IMMEDIATE SOURCE:
; LIBRARY: cDNA library in pBluescript sk (-)
; CLONE: pTAS1
; NAME/KEY: CDS
; LOCATION: 162..2559
; US-09-196-390-5

alignment_scores:
Quality: 1182.50 Length: 825
Ratio: 2.571 Gaps: 24
Percent Similarity: 55.758 Percent Identity: 36.727

alignment_block:
US-09-674-824-2 x US-09-196-390-5 ..
Align seg 1/1 to: US-09-196-390-5 from: 1 to: 2825
```


7 GlyAlaGlyCysLeuAla.....ProSerValArgLeuArgAl 19
||||| |||
267 GGGCGCCGGAGTGTGACTGGCCGCTGGCCGCCGCACGCGAGCTCG 316

19 aAspProAlaThrAlaAla..... 25
||| |||:::|||||
317 CGAGCAGCTGTGGCGCGCTCGCCCGCGGAAGACGCGGGATGG 366

26 ..ArgLaserAlaCysValValArgAlaArgLeuArgAlaArg 41
|||:::||| ||| ::||| |||
367 ACAGACCAGCGCGCTCGTGAGCGACCCCGCGCACTCCGCGGTGGCGCC 416

42 GlyArgTyrValAlaGlu.....LeuSerArgGly 51
::: ||||| |||
417 GCCACCAAGGTCGCGAGCGAGGATCCCGTCAAGACGCTCGACCGCGA 466

51 u.....GlyProAlaAlaArgProAlaGlnGlnGlnL 63
|||:::||| |||:::|||
467 CGCCGCGGAAGCGCGGCCCTCCCGCGCGCAGCGAGCAGCGCG 516

63 euAlaProProLeuValProGlyPheLeuAlaProProPro 76
|||||| ::|||::|||
517 CCGCTCCGCGAGTAGAACGCGCATG.....CCGGTGAAACGCG 554

77Pr 77

555 GAGAACAATCTACCGCGCGCGCGCGACTAAGACAGCGGCTGCC 604

77 oAlaProAlaGlnSerProAlaPro...ThrGlnProProLeuProAspA 93
|||:::||| ||| |||||
605 CACGCGCCAGCACGCGGCCCATCCGTCGACAGCAACAGACACCGGTGA 654

93 laGlyValGlyGluLeu.....AlaProAspleuLeuGlu 105
||| ::||| |||
655 ACGGTGAAAACAAGCTAACGTCGCGCTCGCCGCCAGCAGCATGCCGAG 704

106 GlyIleAlaGluAspSerIleAspSerIleVal..... 117
::|||::||| |||::|||
705 CGCGCGGCTTCGGATTCCGCACTACCATTTCCATCAGCGACAAAGCGCC 754

117 117

755 GGAGTCCGTTGTCAGCTGAGAAGACCGCGCTGTCGCGCTCAAATT 804

118AlaAlaSerGluInAspSerGluIleMetAspAlaAsnGlu 131
::|||::||| |||::|||
805 TCGAGTCTCGGCGCTCTCTCCGCGGTCTGACACTGTCAAGCAGCTGAA 854

132 Gln.....ProGlnAlaLy 136
||| |||::|||
855 CAAGAAGCTCAAGAGGTTGCGGTCTGTCGAAGAGCTCCAAGCCAAA 904

136 sVal..... 137
|||
905 GGCTCTTTTCGCGCTGCAGCCCCCGCTGTACAAAGAGACCTTTGGGATT 954

137 137

955 TCAGAAATACATTGGTTTCGAGGAGCCCGTGGAGCCAAAGGATGATGC 1004

137 137

1005 CGGCTGTCCAGATGATGCGGGCTCTTTGAACACACCAACAGATCAGA 1054

138ThrArgSerIleValPheValThr 146
::|||::||| ::|||::|||
1055 CTCGGACCTTTGGCAGGGGAATGTATGAACGTGTCTCGTGGCTG 1104

146 lyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGly 162
::|||::||| |||::||| |||::|||
1105 CTGAGTGTCTCCCTGGTGCAACACAGTGGTCTGGAGAGTTGTCGGGT 1154

163 SerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValme 179

1155 GCTCTGCCCAAGGCTTTGGCAAGAGAGACATCGTGTATTGTTGGTTGGT 1204
1196 yThrAlaLysHisIleLeuValProCysPheGlyGlySerHisGluVal 212
1243 GAGTCGGAATACTACAAG.....GCTGCTGCAGGATGATGCAAGTG 1286
213 ThrPhePheHisGluTyrArgAspAsnValAspTrpValPheValAspH1 229
1287 AATATTTCATGCTTATATCATCGATGGAGTTGATTTGTGTTCATTGACGC 1336
229 sProSertyr...HisArgProGlySerLeuTyrGlyAspAsnPheGlyA 245
1337 TCCTCTCTCCGACACCGCTCAGGAAGACATTAATGAG.....GGCA 1377
245 laPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystYrAlaAlaCys 261
1378 GCAGACAGGAATATAGAAGCGCATGATTTGTCTCGAAGCGCGCTGTT 1427
262 GluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln...As 277
1428 GAGGTTCATCGCACGTTCATGCGGGCGGTGCCCTTATGGGATGGAAA 1477
277 nCysMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuL 294
1478 TCTGTGTGTTATTCAAATGATGGCACAGGCACTCCTGCCGTGCTATC 1527
294 euAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThr 310
1528 TGAAAGCATATTACAGGAGACCATGTTGTGTCAGTACACTCGTCCATT 1577
311 LeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTy 327
1578 ATGTTGTATACATACATCGTCTACACAGGGCGGTGGCCCTTAGATGAAT 1627
327 rProAspLeuGlyLeuProGluTyrTrpTyrGlyAlaLeuGluTrpValP 344
1628 CCCGTTACCCGAGTTGCTTGAGCACTACCTG..... 1658
344 heProGluTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsn 360
1659GAACACTTCAGACTGTACAGCCCGTGGGTGTGTGAACACGCGAAC 1703
361 PheLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGI 377
1704 TACTTCGCGCGCGCCCTGAGATGCGCGACACAGGTGTCGTGGTGAGCCC 1753
377 nGlyTyrSerTrpGluValThrThrAlaGluGlyGlyGlnGlyLeuAsnG 394
1754 CGGTTACCTGTGGGAGCTGAAGACGGTGGAGGGCGGTGGGGGCTTCACG 1803
394 luLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAlaSnGlyIle 410
1804 ACATCATACGCGAAGACACTGGAAGACCCCGCGGCACTGCTCAACGCAATC 1853
411 AspIleAsnAspTrpAsnProThrThrAspLysCysLeuProHis..... 425
1854 GACAACATGGAGTGGAAACCCCGAGGTGGACGCCACCTCAAGTCGACCGG 1903
426HisTyrSerValAspLeu...SerGlyLysAlaLysCysL 439
1904 CTACACCAACTTCCTCCCTGAGGACGCTGGACTCTCCGCAAGCGCGCATGCA 1953
439 ysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValPro 455
1954 AGGAGGCCCTGCACGGCGACTGGGCTCGAGGTCCGCCGACGCTGGCGG 2003
456 LeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuII 472

```

2004 CTGCTCGCTTTCATCGCGCGCTGGAGCGGAGAAAGCGGTGGAGATCAT 2053
472 eLysMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetL 489
2054 CGCGGAGCCATCGCTGGATCGTGAGCCAGGAGGTGACGTGTGTATGC 2103
489 euGlySerGlyAspProIlePheGluTrpMetArgSerThrGluSer 505
2104 TGGCAGCGGCGCCAGCAGCTGGAGAGCATGCTGCAGCAGCTTCGAGCGG 2153
506 SerTyrLysAspLysPheArgGlyTrpValGlyPheSerValProValSe 522
2154 GAGCACCAGCAAGGTGCGGGTGGGTGGGTTCCTCGTGGCGCTGCG 2203
522 rHiArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheG 539
2204 GCACCGGATCAGCGGGGGGCGGAGCGCTCTCATGCCCTCCCGGTTCG 2253
539 luProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValPro 555
2254 AGCGTGGCGGCTGACACAGCTCTACGCCATGGCTACGGCAGCGTCCCG 2303
556 ValValHisGlyThrGlyLeuArgAspThrValGluThrPheAsnPr 572
2304 GTCGTGACGCGCTCGCGGCGCTCAGGAGACACCGTGGCGGTTCGACCC 2353
572 oPheGlyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuT 589
2354 CTTC.....ACCCTCTCGGCTCGGTGGAGCTTCACCGCGCGCG 2394
589 hrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArg 605
2395 AGCGCACAAAGTATGATCGAGCGCTCGGCGCTGCTCCCGCACCTACCGA 2444
606 GluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAS 622
2445 GACTTCAAGGAGAGGTGAGGGGCGCTCCAGAGCGCGCATGTGCGCAGGA 2494
622 phiThrTrpAspHisAla.ProSerSerThrSerArgSerSerGly 638
2495 CTTCAGCTGGGAGCAGCGCGGCAAGCTCTACGAGGACGCTCCTCGTCAAG 2544
639 ProSerTrpThrAsn.....ProThrSe 646
2545 CCAAGTACAGTGGTGAACGCTAGCTGCTAGCGCTCCAGCGCGCGCATGC 2594
646 rCysArgArgGlyLeuGly.....ArgSerLysCysGlu...S 658
2595 GTGCATGACAGGTGAAGTGCATTCGCGCAGCAGCAAGTGCATGAG 2644
658 exProSerAlaLeuLysThrSer 665
2645 CGCGCGCATCCCGAAGTACAGT 2667

```

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq.US-08-941-445A-8

seq_documentation_block:
: Sequence 8, Application US/08941445A
: Patent No. 6107060

GENERAL INFORMATION:
: APPLICANT: Keeling, Peter
: APPLICANT: Guan, Hanping
: TITLE OF INVENTION: Search Encapsulation
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:

ADDRESSEE: GreenLee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Circle
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2007 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2007
: US-08-941-445A-8

```

alignment_scores:
Quality: 1076.00 Length: 693
Ratio: 2.485 Gaps: 18
Percent Similarity: 62.482 Percent Identity: 37.951

alignment_block:

US-09-674-824-2 x US-08-941-445A-8 ..

Align seg 1/1 to: US-08-941-445A-8 from: 1 to: 2007

```

5 GlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAspPr 21
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12 GCCGGGGGGAAGGAGCGCGCGGAGAGGAGCGCGGAGCGCGCCAG.. 59
21 oAlaThrAlaAlaArgAlaSerAlaCys.ValValArgAlaArgLeuArg 37
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 ....GTTGCCCGCGCTCGGCGCAATGGGTCTCCAAACGAGGAGGATCCT 105
38 ArgLeuAlaArgGlyArgTyr.....ValAlaGluLeuSerArgG1 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 CTTCAGCGCGTCCGCGCGGTACGGCTCCGCGAGCGGGAACACGCGCAGGAC 155
51 uGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 CGGCGCGCGGTCC.....TCCAGACGCGCGCATTTGCGGACGTTGAGA 199
66 ..ProLeuValProGlyPheLeuAlaProProProProAlaProAlaGln 81
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 TCGTTGAGATCAAGTCCATCGTCGCGCGCGCGGAGCAGCATAGTGAAG 249
82 SerProAlaProThrGlnProLeuProAspAlaGlyValGlyGluLe 98
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
250 TTCCAGGG.....CGCGGCTACAGATGATCTTCCCTCTGGGACAT 293
98 uAlaProAspLeuLeu..... 104
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 AGCAGCGAGAGTGTCTCTCCAGCGCGCGGAGCGGAGCATGATGATCGCTG 343
105 .....GluGlyLeuAlaGluAspSerIleAspSerIle 115
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

344 CGTTGACGGAGATTCAAATGGAAATTTGCCACCTCTACAGTTCTGAGCCATTGA 399

116 IleValAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluG1 132

394 GTACAGGAGCCCACTTGGGATTTCAAGAAATACATCGGTTTTCAGCAGACC 443

132 nProGlnAlaLys..... 136

444 TGACGAAGCCGAAGGATGATTCACAGGCTTGGTCAGATGATGCTGTTCTT 493

137ValThrArg 139

494 TTGAACATTATGGACAATGATTCGGGCTTTTGGGGGAGAAATGTTATG 543

140 SerIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyG1 156

544 AACGTGATCGTGGTCTCTCAATGTTTCTCATCGTGGTCAAAACAGGTGG 593

156 YLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyH 171

594 TCTTGGAGATGTGTGGGAGCTTTACCCNAAGCTTTACCGAGAAGAGGAC 643

173 isArgValMetValValMetProArgTyrLeuAsnGlySerSerAspLys 189

644 ATCGCTGTTATGGTGTGGTACCAAGTAT.....GGG 675

190 AsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPh 206

676 GACTATGTGGAAGCTTTGATATGGATCCGGAATACTACAAGATGCG 725

206 eGlyGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValA 223

726 AGCAGAGGACCTAGAAAGTAACTATTTCATGCATTTTATGATGAGTCG 775

223 spTrpValPheValAspHisProSerTyrHisArgProGlySerLeuTyr 239

776 ACTTTGTCTTATGATGCTCTTCCGGCACCGTCAAGATGACATATAT 825

240 GlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLe 256

826 GGG.....GGAGTAGGACGAGAAATCATGAAGCGCATGATTTGTT 866

256 uCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrI 273

867 TTCGAAGTTGCTGTGTGAGTTCTTGGCAGCTTCCATGCGGTGGTGTGT 916

273 letYrGlyGln...AsnCysMetPheValValAsnAspTrpHisAlaSer 288

917 GCTACGGAGATGAAATTTGGTGTTTCATTCGCATGAATTCGCACACTGCA 966

289 LeuValProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrAr 305

967 CTCTCGCTGTTTATCTGAAGGCATATTACAGACACCATGGTGTAAATGCA 1016

305 gaspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValG 322

1017 GTACACTCGCTCCGTCCTCGTCATACATACATCGGCACACCGCCCGCTG 1066

322 LuProLaserThrTyrProAspLeuGlyLeuProProGluTrpTyrGly 338

1067 GTCTCTGTACATGAATTCGCTACATGGAATTCGTGAACACTAACCTTCAA 1116

339 AlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLy 355

1117 CATTTCGAGCTGACGATCCCGTC.....GG 1142

355 sgGlyGluAlaValAsnPheLeuLysGlyAlaValVal...ThrAlaAspA 371

1143 TGGCCGAGCAGCCAAACATCTTTGGCGCTGTGTCTCAAGATGCGACACC 1192

371 rgIleValThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGlu 387

1193 GGGTGGTGACTGTACGCCGGCTACCTGTGGAGCTGAAGACATGGTAA 1242

388 GlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgTlysSerValLeuAs 404
1243 GCGCGCTGGGGCTCCACGACATCATCGTTCTTAACGACTCGAAGACAA 1252
404 nGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspL 421
1293 TGCATTCGTGAACGATCGACACACGAGGTGGAACCCCAAGGTGGACG 1342
421 yScyLeuProHis.....HisTyrSerValAspLeu... 432
1343 TGCACCTGGCGTGGACGGCTACACCACTACTCCTCGAGACACTCGAC 1392
433 SerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuPr 449
1393 GCTGGAAGCGCAGTGGAAAGCGGCCCTCAGCGGGACGCTGGGCCCTGGA 1442
449 oValArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrG 466
1443 AGTGCAGCAGCAGTGCCTCGCTTCGGCTTCATCGGGCTCTGGATCGAC 1492
466 InLysGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGlu 482
1493 AGAAGGCGTGCACATCATCGGGGACGCGATCGCTGGATCGCGGGCAG 1542
483 AspValGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTr 499
1543 GACGTGCAGCTGGTGTGCGGACCGCCACCTGACCTGGAACGAAT 1592
499 pMetArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValG 516
1593 GCTGCAGCACTTGGAGCGGGACATCCCAAGGTGCGGGGTGGGTGCG 1642
516 lyPheSerValProValSerHisArgIleThrAlaGlyCysAspIleLeu 532
1643 GGTTCGTGCTCAATGTGTCATCGCATCGCCGCGGGCGCCGCGTCTG 1692
533 LeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMe 549
1693 GTGATTCCTCCCGTTCGCGCGC...GGGCTGAACACAGCTCTACGCGAT 1739
549 tGlnTyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspT 566
1740 GCATATCGCACCGCTCCCTGTGTGTGCACCGCGTGGCGGGCTCAGGGACA 1789
566 hrValGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGly 582
1790 CCGTGGCGCGGTTCGACCCGTT.....GGCGAGCGGGCTCGGG 1830
583 TrpAlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgTh 599
1831 TGGACTTTTGAACGCGCGAGGCCAACAGTGATGTCAGGTGCTCAGGCCA 1880
599 rAlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetL 616
1881 CTGCCCTGCACAGTACCGAAACTACGAGGAGAGCTGGAAGAGTCTCCAGG 1930
616 yArgGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerTh 632
1931 CGCGGGCATGTCGCAGAACCTCAGCTGGGACCACCGCGGCTGAGCTCTAC 1980
632 rSerArgSerSerGlyProSer 640
1981 GAGGACGCTCTTGTCAAGTACCACT 2005

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-572-951-2

```
seq_documentation_block:
; Sequence 2, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
```

```

; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: MODIFICATION OF STARCH
; TITLE OF INVENTION: SYNTHESIS IN PLANTS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: Pillsbury Madison & Sutro LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,951
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,602
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/263,921
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul N. Kokulis
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222957/1.02.15C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2085 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-572-951-2

alignment_scores:
    Quality: 1071.00      Length: 643
    Ratio: 2.619          Gaps: 14
    Percent Similarity: 63.608      Percent Identity: 38.725

alignment_block:
US-09-674-824-2 x US-08-572-951-2      ..
Align seg 1/1 to: US-08-572-951-2 from: 1 to: 2085

68 valProGlyPheLeuAlaProProProProAlaProAlaGlnSerProAl 84
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
23 ATCAAGTCCATCGTCGCCGCCGCCGCCGACGACATAGTAGTTCGCCAGG 72
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
84 aProThrGlnProProLeuProAlaGlyValGlyGluLeuAlaProA 101
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
73 G.....CGCGGCTACAGGATGATCTCTCTCTCTGGGACATAGCGCGG 116
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
101 splLeuLeu.....| | | | | | | | | | | | | | | | | | |
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
117 ACAGTCTCTCTCCAGCCCGAAGCCACTGCATGAATCGCTCGGCTTGAC 166
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
105 .....GluGlyIleAlaGluAspSerIleAspSerIleIleVal 118
   ::::::::::::::| | | | | | | | | | | | | | | | | | |
167 GGAGATTCAATGGAATTGCACCTCTCTACAGTTGAGCCATTAGTACAGA 216
   ::::::::::::::| | | | | | | | | | | | | | | | | | |

```

```

118 aAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnA 135
   |||::: ::::: ::::: ::::: |||:::
217 GGCCACTTGGGATTTCAGAAATACATCGGTTTGTGACGACCTGACGAAG 266
   |||::: ::::: ::::: ::::: |||:::
135 lAlys.....| | | | | | | | | | | | | | | | | | |
   |||:::
267 CGAAGGATGATTCCAGGGTTGGTGCAGATGATGCTGGTTCTTTTGAACAT 316
   |||:::
137 .....| | | | | | | | | | | | | | | | | | |
   :::::
317 TATGGGACAATGATTCTGGGCCCTTTGTGGGAGAAATGTTATGAACGTGAT 366
   :::::
142 lPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyA 159
   : |||::: ::::: |||::: |||::: |||::: |||::: |||:::
367 CGTGGTGGCTGCTGAATGTTCTCATGGTGCATAAAGAGGTGGTCTTTGGAG 416
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
159 spValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgVal 175
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
417 ATGTTGGGGAGCTTTACCCAAGCTTTACCGAGAGAGGACATCGTGT 466
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
176 MetValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAl 192
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
467 ATGTTTGTGTACCAAGGTAT.....| | | | | | | | | | | | |
   :::::
192 aLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlyS 209
   : |||::: ::::: |||::: |||::: |||::: |||::: |||:::
499 GGAAGCCTTTGATATGGAATCCGGAATACTACAAAGCTGCAGGACAGG 548
   ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
209 erHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrpVal 225
   : |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
549 ACCTAGAAGTGAATTTCCATGCATTTATTTGATGGAGTCGACTTTGTG 598
   ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
226 PheValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAspAs 242
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
599 TTCATTGATGCTCTTCCGCGCACCGTCAAGATGACATATATGG..... 643
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
242 nPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystyrA 259
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
644 ....GGAAGTAGGCAGGAGAAATCATGAAGCCGATGATTTGTTTTCGAAG 689
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
259 laAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGly 275
   : |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
690 TTGCTGTTGAGTTCCTTGGCACGCTTCATGCGGTGGTGTGTGTACGGA 739
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
276 Gln...AsnCysMetPheValValAsnAspTrpHisAlaSerLeuValPr 291
   ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
740 GATGAAATTTGGTGTTCATTGCTCATGAATTGGCACACATGCACCTCCTGCC 789
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
291 oValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerA 308
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
790 TGTTCATCTGAAGCATATTACAGACCATGGGTAAATGCAGTACACTC 839
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
308 rgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAla 324
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
840 GCTCCGCTCTCTCATATCAATACATCGCCACACAGGCCGCGGTGCTCTGTA 889
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
325 SerThrTyrProAspLeuGlyLeuProGluTrpTyrGlyAlaLeuGlu 341
   ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
890 CATGAATTCCTCGTACATGGACTTCTGTAACACTAACCTTCAACATTTTCA 939
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
341 uTrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyGluA 358
   : |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
940 GCTGTACCATCCCGTC.....| | | | | | | | | | | | |
   :::::
358 laValAsnPheLeuLysGlyAlaValVal...ThrAlaAspArgIleVal 373
   : |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
966 ACGCCAATCTTTGCCCGCTGTGTTCTGAAGATGGCAGACCCGGGTGGTG 1015
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
374 ThrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGlu 390
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1016 ACTGTACCCCGGCTACCTGTGGGAGCTGAAGACAGTGAAGGCGGCTG 1065
   ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
390 nGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleV 407
   ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::

```


586 GCNGAYGCGNCCNGCNACNGAYGCGNCGNCGNCGNCCNTAYAYMG 635
126 eMeTAspAlaAsnGluGlnProGlnAla...LysValThrArgSerIleV 142
636 NGARGAYAAAYGARGCNGCGNCCNTYNGCGNCCNAAYGTNATGAAYGTNG 685
142 alpheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGly 158
686 TNGTNGTNGCNWSNGARTGYCCNCCNTYTGAAACNGCGNGGNYTNGGN 735
159 AspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgVa 175
736 GAYGTNGTNGCGNCGNYTNCNAARGCNYTNGCNMNGMNGCNCAIMNGT 785
175 lMeTValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrA 192
786 NATGNTGNTNATHCCNMGNTAY.....GNGARTAYG 817
192 laLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGly 208
818 CNGARGCNGMGAYTYTNGGNTNGMNGMNGNTAYAAAGTNGCGNCGNAR 867
209 SerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrpVa 225
868 GAYWSNGARGTNACNTATYTTTCAYWSNTAYATHGAYGGNGTNGAYTYGT 917
225 lPheValAspHisProSerTyr...HisArgProGlySerLeuTyrGlyA 241
918 NTYTGTCNARGCNCNCCNTYTMGNCAVMGNCAVAAYAAAYATHAYGNG. 966
241 spAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCys 257
967GGNGARMGNTNGAYATHYTNAAARMGNATGATHYTNITYTYG 1008
258 TyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTy 274
1009 AARGCNGCGNTNGARGTNCNCCNTGGTAYGCNCCNTCYGGNGNACNGTNTA 1058
274 rGlyGln...AsnCysMetPheValValAsnAspTrpHisAlaSerLeuV 290
1059 YGGNGAYGNAAYTYTNGTNTTATHGCAAYGAYTGCGAYACNGCNGNY 1108
290 alProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAsp 306
1109 TWCCNGTNTAYTNARGCNTAYATYMGNGAYAYGGNYTNATGCARTAY 1158
307 SerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluPr 323
1159 GCNMGWSNGTNTYNGTNATHCAYAAAYATHGCAAYCARGGNGMNGNCC 1208
323 oAlaSerThrTyrProAspLeuGlyLeuProProGluTrpTyrGlyAlaL 340
1209 NGTNGAYGAYTYTGTNAAAYTYGAYTYNCCN..... 1239
340 euGluTrpValPheProGluTrpAlaAlaArgGHisAlaLeuAspLys... 355
1240GARGATYATATHGAYCAYTY 1260
356GlyGluAlaValAsnPheLeuLysGlyAl 365
1261 AARYTNTAYGAYAAAYATHGNGNGNGAYCAYWSNAAAYGTNTTYGCGNCG 1310
365 aValValThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpG 382
1311 NYTNAARACNGCGNGAYMGNTGTNACNCTNWSNAAAYGNTAYATGTGG 1360
382 luValThrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuSerSer 398
1361 ARYTNAARACNSNGARGNGGNTGGGNYTNCAYGAYATHATHAAYCAR 1410
399 ArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTr 415
1411 AAYGAYTGGAARYTNCARGCNATHGTNAAAYGGNATHGAYATGWSNCGARTG 1460

415 pasnProThrThrAspLysCysLeuProHisHisTyrSerValAspAspL 432
1461 GAAYCCNGCNGTNGAYGTN.....CAYTYNCAYSWSNGAYGAYT 1498
432 eu.....SerGlyLysAlaLysCysLys 439
1499 AYACNAAYTAYACNTTYGARACNTYNGAYACNGGNAARMGNCARTGYAAR 1548
440 AlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLe 456
1549 GCNCNTYNCARMGNCARTYNGGNTNCARGTNGMNGCAYGAYGTNCCNYT 1598
456 uileGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleL 473
1599 NATHGGNTTYATHGGNMGNNTNGAYCAYCARAARGGNGTNGAYATHATHG 1648
473 ysMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeu 489
1649 CNGAYGCNATHCAYTGATHGCGNCGNARGAYGTNCARNTYNGTNGATGYN 1698
490 GlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSe 506
1699 GGNACNGNMGNGCNGAYTYNGARGAYATGYTNMNGMNTTYGARWSNGA 1748
506 rTyrLysAspLysPheArgGlyTrpValGlyPheSerValProValSerH 523
1749 RCAYWSNGAYAAARGTNGCNGTGGTGGTNGGNTTYWSNGTNCNCCNYTNC 1798
523 isArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGlu 539
1799 AYMGNTATHACNGCGNCGNCGAYATHYTNATGCCNWSNMGNNTTYGAR 1848
540 ProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVa 556
1849 CCNTGYGNTNAAAYCARYTNTAYGCNATGCNTAYGNGNACNGTNCNGT 1898
556 lValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProp 573
1899 NGTCAYGCGNTGNGGNGNYTNGMNGAYACNGTNGCNCNTTYGAYCCT 1948
573 heGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThr 589
1949 TY.....AAYGAYACNGGNTNGTNGACNTTYGAYMGNGCNGAR 1989
590 ValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgG 606
1990 GCNAAVMGNATGATHGAYGCNTYTNWSNCAYTYGTNACNACNTAYMGNA 2039
606 uHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspH 623
2040 YTAYAAARGARWSNTGGMNGCNGTGYMGNGCNGMNGNATGCGNARGAY 2089
623 lThrTrpAspHisAla 628
2090 TNWSNTGGGAYCAYGCN 2106

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-6

seq_documentation_block:
; Sequence 6, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US

```

; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 30-SEP-1997
; APPLICATION NUMBER: US/08/941,445A
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 453...2282
; US-08-941-445A-6

```

```

alignment_scores:
  Quality: 900.00      Length: 634
  Ratio: 2.261         Gaps: 24
  Percent Similarity: 62.776   Percent Identity: 36.909

```

```
alignment_block:
US-09-674-824-2 x US-08-941-445A-6 ..

```

```
Align seg 1/1 to: US-08-941-445A-6 from: 1 to: 2542
```

```

73 AlaProProProAlaProAlaGlnSerProAlaProThrGlnProPr 89
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
439 GCTAGACAACACCACATGT...CGGCTCTACACAGCTCCAGCTCGCCACC 485
89 oLeuProAspAla..GlyValGlyGlu...LeuAlaProAspLeuLeu 104
||| ||| ||| ||||| ||||| ||||| ||||| |||||
486 TCGGCCACCGGCTTCGGCATCGCGACAGGTGCGCGCGCTCGTGTCT 535
104 u.....Gluclytle.....AlaGluAspSerI 112
| ||||| ||||| ||||| ||||| ||||| ||||| |||||
536 CCGCCACGGGTTTCAGGCGCTCAAGCCCCCAGCCCGCGCGCGGAGC 585
112 leAspSerIleValAlaAlaSerGluGlnAspSerGluIleMetAsp 128
||| ||| ||| ||||| ||||| ||||| ||||| |||||
586 CGACGTCGCTCAGCGTGACGACACCGCGCGGCGG..... 617
129 AlaAsnGluGlnProGlnAlaLysValThrArg..... 139
||| ||| ||| ||||| ||||| ||||| ||||| |||||
618 GCGACGCCCCAAGCAGCAGCGGTGCGGTGACGCGTGGCAGCGGAGTTCC 667
140 .....SerIleValPheValT ..... 145
668 CTCGGTCGTGCTACGCCACCGCGCGCGCATGACGTCGTGTCGTCG 717
145 hrGlyGluAlaAlaProTyTyAlaLysSerGlyGlyLeuGlyAspValcys 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 GCGCCGAGATGGCCCTCGAGCAAGACCGCGCGCTCGGTGACGTCCTC 767

```

```

162 GlySerLeuProIleAlaLeuAlaAArgGlyHisArgValMetValva 178
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
768 GGTGGCTCCCTCCCTCGCATGGTGCAGATGGTGCACAGGGTCTATGGTGT 817
178 lMetProArgTyTyLeuAsnGlySerSerAspLysAsnTyTyAlaLysAla 195
| ||||| ||||| ||||| ||||| ||||| ||||| |||||
818 CTCCTCCCTCGGTAC.....GACCAGTACAAGGACGCTT 849
195 euTyTyThrAlaLysHisIleLysIleProCysPheGlyGlySerHisGlu 211
||| ||| ||| ||||| ||||| ||||| ||||| |||||
850 GGGATACACGCGTGTGGTGCAGATCAAGGTGTCAGACAGGTACAGAGG 899
212 ValThrPhePheHisGluTyTyArgAspAsnValAspTrpValPheValas 228
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
900 GTGAGGTTTTTCCATTCGTACAGCGGTGAGTGCACCGTGTGTTCATCGA 949
228 pHisProSerTyTy.....HisArgProGly...SerLeuT 239
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
950 CCATCCGCTCATCTCTGGAGAAGGTTTGGGAAAGACCGGTGAGAGATCT 999
239 yrGlyAspAsnPheGlyAla...PheGlyAspAsnGlnPheArgTyTyThr 254
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1000 ACGACCTGACACTGGAGTTGATTACAAGACAAACACAGATCGCTTCAGC 1049
255 LeuLeuCystTyTyAlaAlaCysGluAlaProLeuIleLeuGluGlyG 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1050 CTTCTTTGCCAGGACGACCTCGAGGCTCTAGGATCCTAAACCTCAACAA 1099
271 yTyTy.....IleTyTyGlyGlnAsnCysMetPheValVala 283
| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1100 CAACCCATACCTCAAGGAACCTTATGTTGAGGATGTTGTTGTCGTGCA 1149
283 snAspTrpHisAlaSerLeuValProValLeuLeuAlaLysTyTyArg 299
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1150 ACGACTGGCACACTGGGCCACTGGCCAGCTACCTGAAGAAACAACACTACCAG 1199
300 ProTyTyGlyValTyTyArgAspSerArgSerThrLeuValIleHisAsnLe 316
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1200 CCCAATGGCATCTACAGGAATGCAAGGTGCTTCTGTCATCCACCAACAT 1249
316 uAlaHisGlnGlyValGluProAlaSerThrTyTyProAspLeuGlyLeuP 333
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1250 CTCCTACACGAGGCGGTTTCGCTTTTCGAGGATTAACCTGAGCTGAACCTCT 1299
333 roProGluTrpTyTyGlyAlaLeuGluTrpValPheProGluTrpAlaArg 349
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1300 CCGAGAGGTTTCAGGTTCATCTTCGATTTTCATC..... 1331
350 ArgHisAlaLeuAsp.....LysGlyGluAlaValAsnPheLeuLys 363
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1332 ...GACGGGTATGACACGCGCGGTGGAGGCGAGGAAGATCAACTGGATGAA 1378
363 sGlyAlaValValThrAlaAspArgIleValThrValSerGlnGlyTyTyS 380
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1379 GGCCGAATCTCTGGAAGCCGACAGGCTGCTACCGTGAAGCCCTACTACG 1428
380 erTrpGlu...ValThrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeu 395
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1429 CCGAGGAGCTCATCTCCGGCATCGCCAGGGGATGCGAGCTCGACACATC 1478
396 LeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspI 412
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1479 ATG.....CGGCTCACCGGCATCACCGGCATCGTCAACGGCATGACGCT 1522
412 eAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyTy...S 428
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1523 CAGCAGTGGGATCTCTAGCAAGGACAAGTACATCACCGCCAGTACGAGC 1572
428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
||| ||| ||| ||||| ||||| ||||| ||||| |||||
1573 CAACACCGCATCTGAGGCGAAGGCCCTGAACAAGAGGCGGTGTCAGCGC 1622

```

445 GluLeuGlyLeuProValArgGluAspValProLeuIleClyPheIleGI 461
||| ||||||||| :~: :~: |||||||||
1623 GAGCGGGGTCTCCGGTCGACAGAAATCCCACACTCATCGGTTCAATCG 1672
||| ||||||||| :~: :~: |||||||||
461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1673 CAGGCTGGAGAACAGAGGGCCCTGACGTCATGGCGCGGCATCCCG 1722
||| ||||||||| :~: |||||||||
478 luLeuMetArgGluAspValcInPheValMetLeuGlySerGlyAspPro 494
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1723 AGCTCATCGACGAGCGTCCAGATCGTTCTTGSGTAGTGAAGAAG 1772
||| ||||||||| :~: |||||||||
495 llePheGluGlyTrpMetArgSerThrGluSerSerTyrIlysAspLysph 511
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1773 AAGTTCAGAAAGCTGCTCACAGAGCATGGAGAGAAGTATCCGGGCAAGT 1822
||| ||||||||| :~: |||||||||
511 eArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAlaag 528
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1823 GAGGCGGTGGTAAGTTCAACGGCGCGCTGCTCATCTCATCATCGCGC 1872
||| ||||||||| :~: |||||||||
528 lyCyAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1873 GAGCGAGCTGCTCGCGCTCCCCAGCGCTTCGAGGCCCTGTGACACTCATC 1922
||| ||||||||| :~: |||||||||
545 GlnLeuTyrAlaMetGlnTyrGlyThrValProvalValHisGlyThrGl 561
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1923 CAGCTGCAGGGGGTCAGATACAGGAGCCCTGTGTTCGCGCTCCACCG 1972
||| ||||||||| :~: |||||||||
561 yGlyLeuArgaspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578
||||||| :~: ||||||||| ||| :~: :~: |||||||||
1973 TGGGCTCGTGACAGCGTCATC.....G 1995
||| ||||||||| :~: |||||||||
578 luGluGlyThrGlyTrpAlaPheSerProLeuThrValasp..... 591
||| ||||||||| :~: |||||||||
1996 AAGGCAAGACTGGTTCCACATGGCGCTCTCAGCGTCGACTCAAGGTG 2045
||| ||||||||| :~: |||||||||
592LysMetLeuTrpAlaLeuArgThrAlaMe 601
||||||| :~: ||||||||| ||| :~: :~: |||||||||
2046 GTGGAGCAAGCGAGCTGAAGAAGGTGGCGGCACCCTGAAGCGCGCAT 2095
||| ||||||||| :~: |||||||||
601 tSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgg 618
||||||| :~: ||||||||| ||| :~: :~: |||||||||
2096 CAAGGTCTGTC.....GGCAGCGCGCGGTACGAGGAGATGGTCAGGAAC 2139
||| ||||||||| :~: |||||||||
618 lywMetLysAspHisThrTrpAspHisAlaProSerSer..ThrSerAr 634
||| ||||||||| :~: ||||||||| ||| :~: :~: |||||||||
2140 GCATGNACCAAGGACCTCTCTCTGG...AAGGGCCCTGCGAAGAACTGGGAG 2186
||| ||||||||| :~: |||||||||
634 gSerSerGlyProSerThrTrpAsnProThrSerCysArgArgGly 650
||| ||||||||| :~: ||||||||| ||| :~: :~: |||||||||
2187 ATATGTCCTCTGG..GCCTGGGGTGGCGCC...GCAGCGCGCGGGGA 2229

Sequence	Strd Orig	zScore	EScore	Len	! Documentation
/cgn2_6/ptodata/1/pna/US0937_COMB	seq:US-09-345-214-5	+ 2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-5	+ 2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-9	+ 2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US0966_COMB	seq:US-09-654-617-385863	+ 2821.50	2921.77	2.1e-153	2491
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-385863	+ 2821.50	2921.77	2.1e-153	2491
/cgn2_6/ptodata/1/pna/US0931_COMB	seq:US-09-345-214-12	+ 2806.50	2908.38	1.2e-152	2008
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-12	+ 2806.50	2908.38	1.2e-152	2008
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-12	+ 2806.50	2908.38	1.2e-152	2008
/cgn2_6/ptodata/1/pna/US0994_COMB	seq:US-09-402-254-52	+ 2713.50	2812.99	2.4e-147	1749
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-625-406-12	+ 2713.50	2812.97	2.4e-147	1751
/cgn2_6/ptodata/1/pna/US0930_COMB	seq:US-09-077-564-1	+ 2655.50	2748.28	9.8e-144	2992
/cgn2_6/ptodata/1/pna/US096C_COMB	seq:US-09-654-617-451753	+ 2578.00	2670.33	2.1e-139	2491
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-451753	+ 2578.00	2670.33	2.1e-139	2491
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-684-016-451753	+ 2578.00	2670.33	2.1e-139	2491
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-625-406-20	+ 2546.50	2640.25	1.0e-137	1620
/cgn2_6/ptodata/1/pna/US0933_COMB	seq:US-09-345-214-6	+ 2495.00	2587.28	9.1e-135	1528
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-6	+ 2495.00	2587.28	9.1e-135	1528
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-10	+ 2495.00	2587.28	9.1e-135	1528
/cgn2_6/ptodata/1/pna/US06031_COMB	seq:US-60-312-344-3744	+ 2424.50	2512.47	1.3e-130	1818
/cgn2_6/ptodata/1/pna/US0933_COMB	seq:US-09-345-214-11	+ 2035.00	2111.40	2.9e-108	1419
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-11	+ 2036.00	2111.40	2.9e-108	1419
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-604-436-11	+ 2036.00	2111.40	2.9e-108	1419
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-606-304-9	+ 2001.00	2070.78	5.3e-106	2360
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-606-304-3	+ 1731.50	1793.46	1.5e-90	1758
/cgn2_6/ptodata/1/pna/US097A_COMB	seq:US-09-620-392-7351	- 1523.00	1558.46	1.8e-77	1818
/cgn2_6/ptodata/1/pna/US097A_COMB	seq:US-09-702-134-7216	- 1523.00	1558.46	1.8e-77	1818
/cgn2_6/ptodata/1/pna/US095B_COMB	seq:US-09-534-859-735	- 1517.50	1538.90	2.3e-76	8419
/cgn2_6/ptodata/1/pna/US098_COMB	seq:US-09-803-736-735	+ 1517.50	1538.90	2.3e-76	8419
/cgn2_6/ptodata/1/pna/US096C_COMB	seq:US-09-654-617-258645	+ 1404.50	1455.51	9.9e-72	8419
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-258645	+ 1404.50	1455.51	9.9e-72	8419
/cgn2_6/ptodata/1/pna/US06031_COMB	seq:US-60-312-344-395	+ 1397.50	1449.23	2.2e-71	1300
/cgn2_6/ptodata/1/pna/US094_COMB	seq:US-09-402-254-48	+ 1253.50	1394.52	9.2e-63	2423
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-625-406-10	+ 1246.00	1287.95	2.1e-62	2097
/cgn2_6/ptodata/1/pna/US094_COMB	seq:US-09-402-254-50	+ 1239.50	1281.55	4.9e-62	2010
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-20	+ 1239.50	1280.62	5.5e-62	2248
/cgn2_6/ptodata/1/pna/US096C_COMB	seq:US-09-654-617-269739	+ 1223.00	1269.99	2.1e-61	2248
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-269739	+ 1223.00	1269.99	2.1e-61	2248
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-654-616-269739	+ 1223.00	1269.99	2.1e-61	2248
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-606-304-7	+ 1215.00	1253.37	1.8e-60	2793
/cgn2_6/ptodata/1/pna/US0931_COMB	seq:US-09-345-214-15	+ 1212.00	1252.97	1.9e-60	2019
/cgn2_6/ptodata/1/pna/US0971_COMB	seq:US-09-743-980-15	+			

```
524 GAGCTAAAGTAACACAAAGCATGCTGTTGTAACCGCGAAGCTTCTCCT 573
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
574 TATGCAAAAGCTGGGGGCTAGGAGATGTTGTGGTTCAATTGCCAGTTCG 623
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
624 TCTTGCTGCTCGTGGTCACTGATGGTGTGTAATGCCAGATATTAA 673
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
674 ATGGTACCTCCGATAAAGAAATTATGCAAAATGCAATTTTACACAGAAAACAC 723
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
724 ATTCCGATTCACATGCTTTGGCGGTGAACATGAAGTTACCTTCTCCATGA 773
217 uTyrArgAspAsnValAspTTPValPheValAspHisProSerTyrHisA 234
774 GTATAGACATTCAGTTGACTGGGTGTTGTGATCATCCCTCATATCACA 823
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
824 GACCTGGAAATTTATATGAGATAAGTTTGGTGTCTTTGGTGATAATCAG 873
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
874 TTCAGATACACACTCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsna 284
924 TCAATTTGGAGGATATATTATGGACAGAAATTGCATGTTGTGTGCAATG 973
284 sPrThrHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
974 ATTTGGCATGCCAGTCTAGTGGCAGTCTCTTCTGCTGCAAAATATAGACCA 1023
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
1024 TATGGTGTATTAAAGACTCCCGCAGCATTCCTGTGTAATACATAATTAGC 1073
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
1074 ACATCAGGGGTAGAGCCTGCAAGCACATATCTGACCTTGGGTGGCCAC 1123
334 roGluTrpTyrGlyAlaLeuGluTTPValPheProGluTrpAlaArgArg 350
1124 CTGAATGGGTATGAGCTCTGGAGTGGGTATTCCTGGAATGGCGGAGGAG 1173
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
1174 CATGCCCTTGACAAGGGTGAGGCAGTTAATTTTGAAGGTGCAGTTGT 1223
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
1224 GACAGCAGATCGAATCGTACTGTCTAGTAAGGGTTATTCTCTGGAGGTCA 1273
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1274 CAACTGCTGAGGTGACAGGGGCTCAATGAGCTCTTAAGTCCAGAAAG 1323
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
1324 AGTGATTAACCGAATTTGTAATGGAATTTGACATTAATGATTTGGAACCC 1373
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
1374 TGGCAGACAGAAATGATCCCTGCTCATTTATCTGTTGATGACCTCTCTG 1423
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1424 GAAAGGCCAAATCTAAAGGTGCATTCACAGAGGAGCTGGGTTTACCTATA 1473
```

```
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1474 AGCCCTGATGTTCTCTGATTGGCTTTATTGGAAGGTGGATTATCAGAA 1523
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
1524 AGCATTTGATCTTCAACTTATACCATCAGATCTCATCGCGGAAGATG 1573
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500
1574 TTCAATTTGTCATGCTTGGATCTGGTGACCCACAGAGCTTGAAGATTGGATG 1623
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTyrValGlyPh 517
1624 AGATCTACAGATCGATCTTCAAGGATAAATTTCTGTTGGTGGTGGATTT 1673
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeu 534
1674 TAGTGTTCAGTTTCCACCGAATAACTGCCGCTGCCATATATTGTTAA 1723
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1724 TGCCATCCAGATTCGAACCTTGGTCTCAATCAGCTATATGCTATGCAG 1773
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgaspThrVa 567
1774 TATGGCACAGTTCCTGTTGCTCATGCAACTGGGGGCTTAGAGATACCGT 1823
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
1824 GGAACTTCAACCCCTTTCGGTGAGATGGAGACGAGGTACAGGTGGG 1873
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1874 CATTCGCACCCCTAACACACAGAAAAATGTTGTGGACATTCGGAACCTGCA 1923
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1924 ATATCTACATACAGGAACACACAGTCTCTCTGGGAGGGGCTAATGAAGCG 1973
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
1974 AGGATGTCAAAAGACTTCACGTGGGACCATGCCGCTGAACAAATACGAAC 2023
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgG1 650
2024 AAATCTCCAGTGGGCTTCATCGATCGACCTATGTCATGTAATAAAA... 2071
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
2072 .....AGGACCAAAAGTGGTTCCT.....TGAAGAT 2099
667 erSerPheArgGlyPro.....GluGlyTyrProCys 677
2100 CATCAGTTTCATCATCTTATAGTAACTAAATGATGAAAGAAAAACCCCTGT 2149
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694
2150 ACATTTACAT..... 2158
694 pheAlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrA 711
2158 ..... 2158
711 laSerGlyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727
2159 .....GGAAGGCAGACCGGCTATTGGCTCCATTTGCTCCAACGCTGTC 2200
728 AlaGlyTrpLeuThrAlaLysHis 735
2201 TTTGGCTGGCTTGGCTTCGCTCGATGCAC 2224
```

seq_name: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:US-09-743-980-5

```

seq_documentation_block:
: Sequence 5, Application US/09743980
: GENERAL INFORMATION:
: APPLICANT: E. I. du Pont de Nemours and Company
: TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
: FILE REFERENCE: BB-1147-A
: CURRENT APPLICATION NUMBER: US/09/743,980
: CURRENT FILING DATE: 2001-05-14
: PRIOR APPLICATION NUMBER: 060/094,436
: PRIOR FILING DATE: 1998-07-28
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 5
: LENGTH: 2491
: TYPE: DNA
: ORGANISM: Zea mays
US-09-743-980-5

```

alignment_scores:
 Quality: 2826.50 Length: 742
 Ratio: 4.566 Gaps: 9
 Percent Similarity: 83.423 Percent Identity: 73.989

alignment_block:

US-09-674-824-2 x US-09-743-980-5

Align seg 1/1 to: US-09-743-980-5 from: 1 to: 2491

```

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuAlaAs 20
: ::::::::::::::::::::
157 TCGCGCGTGGCGCGCGTGCCTCCTC..... 183
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
: ::::::::::::::::::::
184 .CTGCGCGCGCGCGCGCGCGCGTGGCGCGCGCGCGCGCGCGC 232
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
: ::::::::::::::::::::
233 GCGCGCTCCAGCGCGTGTGCGCGCGCGCGTGTGCGCGCGCGCGCGCG 282
51 GluGlyProAlaAlaArgProGluAlaGlnGlnGlnLeuAlaProProLe 67
: ::::::::::::::::::::
283 GAGGGCGCGCGCGCGCGCGCGTGCACCGCGCGTGTGCGCGCGCGCT 332
67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
: ::::::::::::::::::::
333 CGTGCGCGGTCTCTCGCG...CGCGCGCGCGCGCGCGCGGTGAGCGG 379
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
: ::::::::::::::::::::
380 CATCGACCGCGCGCGCGTGTGCGCGCGCGCGTGGGGACCTCGGT... 426
101 AspleuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
: ::::::::::::::::::::
427 ...CTCGAACCTGAAGGGATTGTGAAGGTTCATCGATAACACAGTAGT 473
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
: ::::::::::::::::::::
474 TGTGGCAAGTGACCAAGATTCTGAGATTGTGGTTGGAAAGGAGCAAGCTC 523
134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
: ::::::::::::::::::::
524 GAGCTAAAGTAACACAAAGCATTTCTTTCTAACCGCGCGAGGCTTCTCCT 573
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
: ::::::::::::::::::::
574 TATGCCAAGTCTGGGGTCTFAGGAGATGTTGTGGTTCAATGCCAGTTGC 623
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184

```

```

624 TCTTGCTGCTGCTGCTACCGTGCATGGTTGTAATGCCAGATATTAA 673
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
: ::::::::::::::::::::
674 ATGTACCTCCGATGAAGAAATATCAAAATGCAATTTACACAGAAAAAC 723
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
: ::::::::::::::::::::
724 ATTCGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATGA 773
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHis 234
: ::::::::::::::::::::
774 GTATGAGATTTCAGTTGACTGGGTGTTGTTGATCATCCCTCATATACA 823
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
: ::::::::::::::::::::
824 GACCTGGAAATTTATATGGAGATAAGTTGGTGCTTTGGTGATAATCAG 873
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
: ::::::::::::::::::::
874 TTCAGATACACACTCCTTGTCTATGCTGCTGATGTGAGGCTCCTTTGATCCT 923
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsn 284
: ::::::::::::::::::::
924 TGAATTCGGAGGATATATTTATGACAGAAATGCAATGTTGTTGTCAGT 973
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
: ::::::::::::::::::::
974 ATTGGCATGCCAGTCTAGTGCAGTCTTCTTGTGTCGCAAAATATAGACA 1023
301 TyrGlyValTyrArgAspSerArgSerThrLeuValLeuIleHisAsnLeu 317
: ::::::::::::::::::::
1024 TATGGTGTTTATAAAGACTCCCGCAGCATTTCTTGTATAACATAATTTAGC 1073
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 334
: ::::::::::::::::::::
1074 ACATCAGGCTGTAGAGCTCTCAAGCACATATCTGACCTTGGGTGGCCAC 1123
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
: ::::::::::::::::::::
1124 CTGAATGGTATGAGGCTCTGGAGTGGGTATTCCTGAAATGGCGGAGAGG 1173
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
: ::::::::::::::::::::
1174 CATGCCCTTGACAGGGTGAGGCGAGTTAAFTTTTGAAGGTGCAGTTGT 1223
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluVal 384
: ::::::::::::::::::::
1224 GACAGCAGATCGAATCGTGACTGTCAAGTAAAGGTTATTCGTGGGAGTCA 1273
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
: ::::::::::::::::::::
1274 CAACCTGCTGAAGGTGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1323
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
: ::::::::::::::::::::
1324 AGTGATTTAAACGGAATTTGAAATGGAATGACATTAATGATTTGAACCC 1373
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
: ::::::::::::::::::::
1374 TGCCACAGACAAATGATCCCTGTCTATTATTCTGTTGATGACCTCTCTG 1423
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
: ::::::::::::::::::::
1424 GAAAGCCCAATGTAAGGTGCAATTCAGAAAGAGCTGGGTTTACCTATA 1473
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
: ::::::::::::::::::::
1474 AGGCCTGATGTTCCCTCTGATTGGCTTTATTTGAAGGTTGATTTATCAGAA 1523
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
: ::::::::::::::::::::

```

1524 AGCATTGATCTCATTCAACTTATCATACCAGATCTCATGCGGGAAGATG 1573

484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||
1574 TTCATTTGTCATGCTTGGATCTGGTGACCCAGACGCTTCAAGATTGGATG 1623
|||||
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
|||||
1624 AGATACACAGATCGATCTTCAAGGATAAATTTCTGTTGGATGGTTGGATT 1673
|||||
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
|||||
1674 TAGTGTTCACAGTTCCACCGAATAACTGCGGCTCGCATATATTCTTAA 1723
|||||
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||
1724 TGGCATCCAGATTCCGAACCTTGTGGTCTCANTCAGCTATATGCTATGCAG 1773
|||||
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||
1774 TATGGCACAGTCTCTGTGTCCATGCAACTGGGGGCTTAGAGATACCGT 1823
|||||
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
|||||
1824 GGAGAACTTCAACCCCTTCGCTGAGAAATGGAGACAGGGTACAGGTGGG 1873
|||||
584 laPheSerProLeuThrValIAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||
1874 CATTCGACCCCTAACACACAGAAACATGTTGTGGACATTGCCAAGTGA 1923
|||||
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
.: :
1924 ATATCTACATACAGGGAACACAAGTCTCTCTGGGAAGGGCTAATGAAGCG 1973
|||||
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
.: :
1974 AGGCATGTCAAAAGACTTCACGTGGGACCATCGCTGAAACAATACGAAC 2023
.: :
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
.: :
2024 AAATCTCCAGTGGGCTTCATCGATCGACCCATGTCATGATAAAAAA.. 2071
.: :
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
|||||
2072AGACCAAAGTGTGGTTCCT.....TGAAGAT 2099
.: :
667 eSerPheArgGlyPro.....GluGlyTyrProCys 677
|||||
2100 CATCATGTTTCATCTCTTAGTACCTAATATGATGAAGAAACCCCTGT 2149
.: :
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694
|||||
2150 ACATTACAT..... 2158
.: :
694 pPheAlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrA 711
.: :
2158 2158
.: :
711 laSerGlyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727
|||||
2159GGAAGCAGACCGGCTATTGGCTCCATTGCTCCAAGCTGCG 2200
.: :
728 AlaGlyTrpLeuThrAlaLysHis 735
|||||
2201 TTGTGGCTGGCTTGCCTCGATGAC 2224
|||||

seq name: /cqn2 6/ptodata/1/pna/US6009 COMB.seq:US-60-094-436-9

```
seq_documentation_block:
; Sequence 9, Application US/60094436
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
```

```

; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Modification of Starch Biosynthetic Enzyme Gene
; TITLE OF INVENTION: Expression to Produce Starches in Grain Crops
; FILE REFERENCE: BB-1147-PI
; CURRENT APPLICATION NUMBER: US/60/094,436
; CURRENT FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft word Version 7.0A
; SEQ ID NO 9
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Zea mays
; US-60-094-436-9

```

```

alignment_scores:
  Quality: 2826.50      Length: 742
  Ratio: 4.566          Gaps: 9
  Percent Similarity: 83.423  Percent Identity: 73.989

alignment_block:
US-09-674-824-2 x US-60-094-436-9 ..

Align seg 1/1 to: US-60-094-436-9 from: 1 to: 2491

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
: : : : : : : : : : : : : : : : : : : : : : : : : :
157 TCGGCGCTGGCGCGCGTGCCTCCTC..... 183

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
||| ||||| : : : : : |||||
184 .CTCGGCGGGGCGCGCTGGCGCGCGCGCTGGCGGACCGGCGCGCGCGC 232

35 .ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
: : : : : : : : : : : : : : : : : : : : : : : : : :
233 GCGGCTCCAGCGCGTCTCGCGCGCGCGTGCCTGCGGAGCTGACGAG 282

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProle 67
||||| ||||| ||||| ||||| ||||| ||||| |||||
283 GAGGGCGCGCGCGCGCGCGCTGCCACCGCGCTGCTGGCGCGCGCGCT 332

67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
||| : : : : : ||||| : : : : :
333 CGTGCCGGCTCTCCTCG...CCGCGCGCGGAGGCCACCGGTGAGCGG 379

84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|| |||| : : : : : : : : : : : : : : : : : : : :
380 CATCGACGCGCGCGCGTGCCTGCCGACCGCGCTGGGGGACCTCGGT... 426

101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
||| ||||| ||||| ||||| ||||| ||||| |||||
427 ...CTCGAACCTGAAGGATTGCTGAAGTTCCATCGATAACACAGTAGT 473

117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
||| ||||| ||||| : : : : : ||||| : : : : :

```

201 lIeYsIleProCysPheGlySerHisGluValThrPhePheHisG1 217
724 ATTCGGATTCCATGCTTTGGCGGTCAACATGAAGTTACCTTCTTCCATGA 773
217 uTyArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
774 GTATAGAGATTCAAGTTGACTGGGTGTTTGTGTGATCATCCCTCATATACA 823
234 rProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
824 GACCTGGAAATTTATATGGAGATAGTTGGTCTTTGGTGTGTAATCAG 873
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
874 TTCAGATACACACTCTTGTATGCTGCATGTCAGGCTCCTTTGATCCT 923
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
924 TCAATTGGGAGGATATATTTATGGACAGAATTGCATGTTTGTGTCATG 973
284 sPTripHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
974 ATTCGCATGCCAGTCTAGTCCAGTCTCTTGTCTGCAAAATATAGACCA 1023
301 TyrGlyValTyrArgSerArgSerThrLeuValIleHisAsnLeuAl 317
1024 TATCGTCTTTATAAAGACTCCCGCAGCATCTTGTATAATACATAATTTCAG 1073
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
1074 ACATCAGGGTGTAGAGGCTCAGACCATATCTTGCCTTGGGTGGCCAC 1123
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
1124 CTGAATGCTATGGAGTCTGAGTGGGTATTCCTGTAATGGCGAGGAGG 1173
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
1174 CATGCCCTTGACAAAGGTGAGGAGGTAAATTTTGAAGGTGACGTTGT 1223
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
1224 GACAGCAGATCGAATCGTGAATCTCAGTAAGGGTTATTCGTGGAGGTCA 1273
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1274 CAATGCTGAAGTGGACAGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1323
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
1324 AGTGATTAACCGAATTTGAAATGGAATTGACATTTAATGATTGGAACCC 1373
417 oThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
1374 TGCCACAGACAAATGATCCCTGTCTATTTCTGTTGATGACCTCTCTG 1423
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1424 GAAAGGCCAAATGTAAGGTGATTCAGAAAGGAGCTGGGTTCACCTATA 1473
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1474 AGCCCTGATGTTCTCTGATGGCTTTATTTGGAAGGTTGGATTATCAGAA 1523
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
1524 AGGCATTGATCTCATTCAACTTATCATACAGATCTCATCGCGGAAGATG 1573
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
1574 TTCATTTCTGATGCTGATGCTGGTGACCCAGAGCTTCAAGATTGGATG 1623
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517

1624 AGATCTACAGAGTCGATCTTCAAGGATAAAATTTCTGGATGGGTGGATT 1673
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1674 TAGTGTTCAGATTTCACCGAATAACTGCGGCTGCGATATATTGTAA 1723
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1724 TGGCATCCAGATTGGAACCTTGTGCTCAATCAGCTATATGCTATCAG 1773
551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567
1774 TATGGCACAGTTCTGTGTCCATGCAACTGGGGGCTTATAGATACCGT 1823
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584
1824 GGAGAACTTCAACCTTTCGCTGAGAATGGAGACAGGTTACAGGTGGG 1873
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1874 CATTCGACCCCTAACACACAGAAACATGTTGTGGACATTGCGAACTGCA 1923
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1924 ATATCTACATACAGGAACACAAAGTCTCTCTGGGAAGGCTAATGAAGC 1973
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
1974 AGCATGTCAAAAGACTTCACGTGGGACCATGCCGTGAACAATACGAAC 2023
634 ArgSerSerSerGlyProSerTrpTrpTrpTrpTrpTrpTrpTrpTrp 650
2024 AAATCTTCAGTGGGCTTCATCGACGCCCTATGCTCATGTGTAATAAAA 2071
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
2072AGACCAAGTGGTGGTTCCTTGAAGAT 2099
667 erSerPheArgGlyPro.....GluGlyTyrProCys 677
2100 CATCATTCATCATCTCTAGTAAGCTAAATGATGAAGAAAAACCCCTGT 2149
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694
2150 ACATTACAT 2158
694 pPheAlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrA 711
2158 2158
711 laSerGlyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727
2159GGAAGGCAGACCGCTATTGGCTCCATTGCTCCCAACGCTCTGC 2200
728 AlaGlyTrpLeuThrAlaLysHis 735
2201 TTTGGCTGGCTTGCCTCGATGCAC 2224
seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-654-617-385863
seq_documentation_block:
; Sequence 385863, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 385863
; LENGTH: 2606

```
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-654-617-385863
```

```
alignment_scores:
  Quality: 2821.50      Length: 726
  Ratio: 4.522          Gaps: 18
  Percent Similarity: 85.950  Percent Identity: 76.309
```

```
alignment_block:
US-09-674-824-2 x us-09-654-617-385863 ..
```

```
Align seg 1/1 to: US-09-654-617-385863 from: 1 to: 2606
```

```

2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CGGGGATGGGATCGGGCGCGCTGCTGGTGGCGCGAGGTGAGG... 198
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 .....CGGGGAGG.....AGGT 211
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluSerArg 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 TCGGCTCCAGCGGTGCGGAGG...CGGTGCTGGCGGAGCTGAGCAGG 258
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
259 GACGGTGGGTGGGCG.....CAGGGCGCGTGGCACCGCGGCC 296
66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 GCTGGTGAAGCAGCGGCTCGCGACCTTCTCGTGGCGCGTGGAGCG 346
76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 CACCGCGCCACGACGTCGCGCGCGCGCGCGCGACCGCGCGCGTGG 396
91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyI 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
397 CCGGACTCCGGCGTGGGAGATCGAGCCCGAT.....CTAGAGGTCT 440
107 eaGluAspSerIleAspSerIleValAlaAlaSerGluGlnAsps 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
441 CACAGAAGATTCATCGACAAACAAATTTTGTGGCTAGTGCAGCAGG 490
124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArg 140
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
491 CTGAGATCATGGATGTGAAGGAGCAAGCTCAAGCTAAAGTACACGCC 540
141 ileValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGly 157
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 GTTGTCTTTGTAAACCGGTCAAGCTTCTCTATGCAAAAGTCAGGTG 590
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyH 174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
591 AGGAGATGTTGTGGTTCACTGCCAAATTCCTCTCTGCTGCTGTCAT 640
174 rGValMetValValMetProArgTyrLeuAsnGlySerSerAspLys 190
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
641 GTGTGATGTTGTATGCGGAGATACATGAACGGCGCTTGACAAAAT 690
191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPhe 207
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
691 TTTGCAACGCATTTTACACTCAGAAGCACATTAAGATTCCATGCTT 740
207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnVal 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
741 CGGAGAACATGAAGTACTTTTTTTTTCACGAGTATAGGATTCGTG 790
224 rPValPheValAspHisProSerTyrHisArgProGlySerLeuTyr 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

791 GGGTGTTTGTGATCATCCCTCATATCATAGACCTGGAAATTTGTATG 840
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeu 257
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 GATAATTTTGGTGCCTTTGGCGATAATCAGTTCAGATACACACTCTG 890
257 sTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyr 274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
891 CTATGCGCGGTGTGAAGCCCCCATTAATCTTGAACCTGGGAGGATA 940
274 yrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeu 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
941 ATGGACAGAAATGCATGTTTGTGTGAATGATGGCATGCCACTCTGT 990
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAs 307
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
991 CCACTCTCTCTGCTGCAAAATATACCCATATGTTGTTTACAGGGAT 1040
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGlu 324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1041 CCGCAGTGTCTTGTATCATCATATATAGCACATCAGGGTGTGGACC 1090
324 laSerThrTyrProAspLeuGlyLeuProGluTrpTyrGlyAlaLeu 340
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1091 CCACTATATCTCTGACCTGGGATTCACCTGAATGGTATGGAGCATTA 1140
341 GluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLysGly 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1141 GAATGGGTGTTTCCAGAGTGGCAAGCGGCATGCCCTTGACAAGGTG 1190
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleVal 374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1191 GGCAGTCAATTTTAAAGCGCGAGTTGTGACAGCAGATCGAATTTGT 1240
374 hrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGln 390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1241 CTGTACCCAGGGGTATTATGGAGGTGCACAACTGCTGAAGGTGGCAA 1290
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIle 407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1291 GGCCTCAATGAGTCTTAAAGCTCCCGAAGAGTGTATTGAATGGAAT 1340
407 laAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCys 424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1341 AATGGAAATTCACATTAATGATGGAAACCCATCCACAGACAAAGTT 1390
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLys 440
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1391 CTTATCATTTATCTGTTGATGACCTGTCCGAAAGGCCAAAGTGTAA 1440
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeu 457
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1441 GAATTCAGAGGAGCTGGGTTTACCTATAGGCCCGCATGTGCTCTCAT 1490
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLys 474
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1491 TGGCTTTATTGGAAGATTTGACTATCAAAAGGCATTTGATCTAATAA 1540
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeu 490
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1541 TTGCCATTCACAGATCTCATGCGGACAAATATTCAATTCGTCATGCT 1590
491 SerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSery 507
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1591 TCTGGTCACCCAGGTTTGAAGAGTGGATGAGATCCACAGAAATCAGG 1640
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHis 524
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1641 CAGGGATAAATTCGTGCGGTGGGTTAGTATTAGTGTCCAGTTTCCCAC 1690
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGlu 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1691 GAATAACTGCAGGTTGCGATATATATTGTTGATGCCATCCAGATTCCA 1740

```

```

541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValva 557
|||||
1741 TGTGGCCTCAATCAGCTATATGCTATCCATATGGTACAGTGCTGTGTG 1790

557 lHisGlyThrGlyClyLeuArgAspThrValGluThrPheAsnProPheG 574
|||||
1791 TCATGGAACTGGAGGCCCTCAGAGATACAGTGGAGATTTTAAACCCGTTG 1840

574 lAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590
:::
1841 CTGGAAGAAGCAGACAGGGTACAGGGTGGGCATCTCGCCACTAACCAAT 1890

591 Asp_LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607
|||||
1891 GAAAAAATCTCTGGGCATCTGGGATGGCAATTCGACATACAGGGAAC 1940

607 lLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
|||||
1941 ACAAGTCTCTGGGAGGGTCTAATGAGCGAGGCATGTCAAGCGACTTT 1990

624 ThrTrpAspHisAla_ProSerSerThrSerArgSerSerSerGlyProS 640
|||||
1991 ACATGGGACCATGCCGCTCACAGTATGAACAGATCTTCAATGGCCCTT 2040

640 eTrpThrAsnProThrSerCysArg_ArgGlyLeuGlyArgSerLys.. 655
|||||
2041 CATGGATACCAACATATGTCATGTAATGATTGTGAAGGAAGCAGCAATT 2090

656 CysGluSerPro.....SerAlaLeuLysThrS 665
|||||
2091 TCTCCGAGGACCTCAATCTCTGCTCTTCATGAGCGGAATGAAACATT 2140

665 erSerSerSerPhe_ArgGly.....Pr 672
|||||
2141 TGTACACTACATGGAAAGGGAACCACTTATGCAAAAGTTGCAAAACGATCAC 2190

672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
|||||
2191 TCAAGGTACCCTTGTAGS.....CCTGCTACTTGGCCA..... 2224

689 ysAlaCysLeuLeuTrpPhe 695
|||||
2225 .....ATATGGTTC 2233

```

seq_name: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:US-09-684-016-385863

```

seq_documentation_block:
  : Sequence 385863, Application US/09684016
  : GENERAL INFORMATION:
  : APPLICANT: Kovalic, David K.
  : APPLICANT: Liu, Jingdong
  : TITLE OF INVENTION: Annotated Plant Genes
  : FILE REFERENCE: 38-21(15097)D
  : CURRENT APPLICATION NUMBER: US/09/684,016
  : CURRENT FILING DATE: 2000-10-10
  : PRIOR APPLICATION NUMBER: US 09/654,617
  : PRIOR FILING DATE: 2000-09-05
  : NUMBER OF SEQ ID NOS: 463173
  : SEQ ID NO 385863
  : LENGTH: 2806
  : TYPE: DNA
  : ORGANISM: Oryza sativa
  : US-09-684-016-385863

```

alignment_scores:	
Quality:	2821.50
Ratio:	4.522
Percent Similarity:	85.950
Length:	726
Gaps:	18
Percent Identity:	76.309

alignment_block:

US-09-674-824-2 x US-09-684-016-3858663

```

Align seg 1/1 to: US-09-684-016-385863 from: 1 to: 2606
2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgIle 17
||||: |||:||||:||||:||||:||||:||||:||||:||||:||||:
151 GCGGGATGGGATCGGGCGGGTCCCTGTGTGGCCCGCAGAGTGAAG... 198
17 uArgAlaAspProAlaThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 34
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
199 .....CGGGGAGG.....AGGT 211
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
212 TCGGGCTCCAGCGGGTGGCGAGG...CGGTGGTGGCGAGTGCAGCAG 258
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaPro.... 65
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
259 GACGGTGGTGGCGG.....CAGGGCCGCTGGCACCGCGGCC 296
66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
297 GCTGTGAAGCAGCGCGGTCTCGCAGACCTCTCGTCCGAGTCGACGC 346
76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProLeu 90
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
347 CACCGCGCCGACGCGAGTCGGCGGGCGCGCGCGACCCCGCGCGTGG 396
91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyIle 107
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
397 CGGGACTCCGCGTGGGGAGATCGAGCCCGAT.....CTAGAAGGTCT 440
107 eaGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAsp 124
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
441 CACAGAAGATTCCATCGACAAACAATTTTGTGGCTAGTCAGCAGGAGT 490
124 erGluIleMetAspAlaAsnGluGlnProGlnAlaAlaLysValThrArgSer 140
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
491 CTGAGATCATGGATGTGNAGGAGCAGCTCAAGCTAAGTAACACGAGC 540
141 lleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
541 GTTGCTCTTTGTAACGGTGAAGCTTCTCTTATGCCAAGTCAGGTGGACT 590
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaAargGlyHis 174
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
591 AGGAGATGTTGTGGTTCACCTGCCAATGTCTTGTCTTCGTGGTGCATC 640
174 rgValMetValIleMetProArgTyrLeuAsnGlySerSerAspLysAsn 190
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
641 GTGTGATGGTGTGAATCCCGAGATACATGAACGGGGCTTGAACAAAAT 690
191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
691 TTTGCAAGCATTTTACACTGAGAAGCACATTAAGATTCCATGCTTTGG 740
207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAsp 224
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
741 CGGAGACATGAAGTACTTTTTTTCACGAGTATAGGATCTCTGTCATT 790
224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
791 GGGTGTGTTGTGATCATCCCTCATATCATAGACCTGGAAATTTGTATGA 840
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
841 GATAATTTTGGTCTTTGGCGATAATCAGTTCCAGATACACACTCTGTG 890
257 sTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIle 274
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
891 CTATCGCGGTGTGAAGCCCATTAATTCTTGACCTGGGAGGATATATCT 940
274 yrGlyGlnAsnCysMetPheValAlaAsnAspTrpHisAlaSerLeuVal 290
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

```

Fri Mar 29 09:26:44 2002

```
941 ATGGACAAATGATGTTTCTTGTGAATGATGGCATGCCAGTCTTGTG 990
291 ProValLeuLeuAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
|||||
991 CAGGCTCTTCTGTGTCAAATATAGACCATATGTTGTGTACAGGGATGC 1040
|||||
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
|||||
1041 CGCGAGTGTCTTGTTCATACATAATCTATGACACATCAGGGTGTGAGCGCTG 1090
|||||
324 laSerThrTyrProAspLeuGlyLeuProProGluTyrTyrGlyAlaLeu 340
|||||
1091 CAGATACATATCTGACCTGGGATTCGCACCTGAATGTTGATGAGCATTA 1140
|||||
341 GluTyrValPheProGluTyrAlaArgArgHisAlaLeuAspLysGlyG 357
|||||
1141 GAATGGGTGTTTCCAGAGTGGGCAAGCGGCATGCCCTTGACAAAGGGTGA 1190
|||||
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
|||||
1191 GCGAGTCAATTTTAAAGGCGCAGTGTGACAGCAGATCGAATGTGTA 1240
|||||
374 hrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGln 390
|||||
1241 CTGTGAGCCAGGGGTATTTCATGGGAGGTGCACAACTGCTGAAGGTGGGCAA 1290
|||||
391 GlyLeuAsnGluLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
|||||
1291 GGCCCTCAATGAGCTCTTAAGCTCCCGGAAGAGTGTATGAATGGAAATGT 1340
|||||
407 lAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424
|||||
1341 AAATGGAATGACATTAATGATTGGAACCCATCCACAGACAAGTTCTCC 1390
|||||
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
|||||
1391 CTATCATATTATTCGTGTGATGACCTGTCCGGAAGGCCAAGTGTAAAGCT 1440
|||||
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI 457
|||||
1441 GAATTCAGACAGAGGTGGTGTACCTATTAAGCCCGATGTGCCCTGTAT 1490
|||||
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
|||||
1491 TGGCTTTATTTGGAAGATTGGACTATCAAAAGGCATTTGATCAATTAAC 1540
|||||
474 eAlaIleProGluLeuLeuMetArgGluAspValGlnPheValMetLeuGly 490
|||||
1541 TTGCCATTCAGATCTCATCGCGGACAAATATTCATTCGTCATGCTTGA 1590
|||||
491 SerGlyAspProPheGluGlyTyrMetArgSerThrGluSerSerTy 507
|||||
1591 TCTGGTCACCCAGTTTGAAGGATGGATGAGATCCACAGAAATCAGGGTA 1640
|||||
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524
|||||
1641 CAGGATAAAATTCGTGGATGGGTGGATTAGTGTTCAGATTTCACACC 1690
|||||
524 rGileThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540
|||||
1691 GAATAACTGCAGGTGGGATATATTGTTGATGCCATCCAGATTCGAACCT 1740
|||||
541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557
|||||
1741 TGTGGCTCAATCAGCTATATGCTATGCAATATGTTACAGTGCCTGTGTG 1790
|||||
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574
|||||
1791 TCATGGAACTGGAGGCTTCAGAGATACAGTGGAGAATTTTAACCCGTTG 1840
|||||
574 lyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590
|||||
1841 CTGAGAAAGGAGCAGGGTACAGGGTGGGCATTTCTGCCACTAACCAT 1890
```

```
591 Asp.LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607
|||||
1891 GAAAAAATGCTGTGGGCATTCGGATGCAATTTTCACATACAGGAAC 1940
|||||
607 lLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
|||||
1941 ACAAGTCCCTTTGGGAGGTCTAATGAAGCGAGGCATGTCAAGCCACTTT 1990
|||||
624 ThrTrpAspHisAla.ProSerSerThrSerArgSerSerSerGlyProS 640
|||||
1991 ACATGGGACCATGCCCGCTCACAGTATGAACAGATCTTCAATGGGCCTT 2040
|||||
640 eTrpTrpAsnProThrSerCysArg.ArgGlyLeuGlyArgSerLys.. 655
|||||
2041 CATGGATCAACCATATCTCATGTAATGATTTGAAGGAAGCAGCGAAAT 2090
|||||
656 CysGluSerPro.....SerAlaLeuLysThrS 665
|||||
2091 TCTCCGAGGACCTCAATCTCTCTCTCTTCATGAGCGGAATGAAACTT 2140
|||||
665 erSerSerSerPhe.ArgGly.....Pr 672
|||||
2141 TGTACACTACATGGAAGGCAACCAAGTTATGCAAAAGTTGCAACAGATCAC 2190
|||||
672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
|||||
2191 TCAAGGTTACCTTGTAGG.....CCTGCTACTTGGCCA..... 2224
|||||
689 ysAlaCysLeuLeuTrpPhe 695
|||||
2225 .....ATATGGTTC 2233
|||||
```

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:US-09-345-214-12

seq_documentation_block:

; Sequence 12, Application US/09345214

; GENERAL INFORMATION:

; APPLICANT: Lightner, Jonathan E.

; APPLICANT: Broglie, Karen E.

; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE

; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS

; FILE REFERENCE: BB-1147

; CURRENT APPLICATION NUMBER: US/09/345,214

; CURRENT FILING DATE: 1999-06-30

; EARLIER APPLICATION NUMBER: 060/094,436

; EARLIER FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

; LENGTH: 2008

; TYPE: DNA

; ORGANISM: Zea mays

; US-09-345-214-12

alignment_scores:

Quality: 2806.50 Length: 660

Ratio: 4.717 Gaps: 6

Percent Similarity: 90.152 Percent Identity: 80.909

alignment_block:

US-09-674-824-2 x US-09-345-214-12 ..

Align seg 1/1 to: US-09-345-214-12 from: 1 to: 2008

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20

|||||

14 TCGGCGGTGGGCGCGCGTGCCTCTCTC..... 40

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34

|||||

41 .CTCGCGCGCGCGCGCTGCCCGCGCGCGTCTCGCGCAGCGGCGCGCCGC 89

35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
90 GCGCGCTCCAGCGGTGCTGCGCGCGGTGCTGCGGAGCTGAGCAGG 139
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67
140 GAGGGCCCGCGCGCGCGCTGCCACCGCGCTGCTGGCGCCCGCT 189
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
190 CGTGCCGGCTCTCTCGG...CGCGCGCGAGCCACGGGTGAGCGG 236
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
237 CATCGACCGCGCGCGCGCTGCCAGCGCGGTGGGGACCTCGGT... 283
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIle 117
284 ...CTCGAACCTGAAGGATGCTGAAGGTTCCATCGATAACACAGTAGT 330
117 lAlaIleSerGluGlnAspSerGluIleMetAspAlaGlnGlnProG 134
331 TGTGGCAAGTGACCAAGATTCTCAGATTGCTGTTGGAAGGACGACGTC 380
134 lAlaIleValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
381 GAGCTAAGTACACAAACATGTGCTTTGTAAACCGCGGAAGCTCTCTCT 430
151 TyrAlaIleSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
431 TATGGCAAGTCTGGGGTCTAGGAGATGTTGTGGTTCATTGCCAGTTGC 480
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
481 TCTTGCTGCTGCTGCTACCGGTGATGGTGTATATGCCAGATATTAA 530
184 snGlySerSerAspLysAsnTyrAlaIleAlaLeuTyrThrAlaLysHis 200
531 ATGGTACCTCCGATAAGAAATTATGCAAAATGCATTTACACAGAAACAC 580
201 IleLysIleProCysPheGlyCysSerHisGluValThrPhePheHisG 217
581 ATTCGATTCATGCTTTGGCGGTGACATGAAGTTACCTTCTTCCATGA 630
217 uTyrArgAspAsnValAspTyrPheValAspHisProSerTyrHisA 234
631 GPATAGAGATTTCAGTTGACTGGGTGTTTGTGATCATCCCTCATATCACA 680
234 rGProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
681 GACCTGGAAATTTATATGGAGATAAGTTTGGTGTGTTGGTGATAATCAG 730
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
731 TTCAGATACACACTCTTTGCTATGCTGCATGTGAGGCTCTTTGATCCT 780
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
781 TGAATGGAGGATATATTATGGACAGAAATTCATGTTTGTGTCAATG 830
284 spTrHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
831 ATTGCAATGCCAGTCTAGTGCCAGTCTTCTTGCTGCAAAATATAGACCA 880
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
881 TATGGTGTATTAAAGACTCCCGACGATCTTGTGTAATACATAATTAGC 930
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
931 ACATCAGGCTGTAGACCTTGCAAGCACATATCTGACCTTGGGTGGCCAC 980

334 roGLuTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
 981 CTGAATGGTATGGAGCTCTCGAGTGGGTATTCCTCAATGGCGAGGAGG 1030
 351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
 1031 CATGCCCTTGACAAGGGCTGAGGCAGTAAATTTTGTAAAGGTGCAGTTGT 1080
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
 1081 GACAGCAGATCGAATCGTGACTGTCAGTAAGGGTTATTCTGGTGGAGGTCA 1130
 384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
 1131 CAATCTGCTGAAGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCAGAAAG 1180
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
 1181 AGTGATTAACCGGAATTCGAATGGAATTGACATTAATGATTGGAAACCC 1230
 417 oThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
 1231 TGGCCACAGACAATGATGCCCTGTCATTATCTGTTGATGACCTCTCTG 1280
 434 lYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
 1281 GAAGGCCAAATGTAAGGTGCATTGCAAGAGGAGCTGGGTTTACCTATA 1330
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLys 467
 1331 AGGCCTTGATGTTCCTCTGATTGGCTTTATTGGAAGCTTGATTATCAGAA 1380
 467 sGlyIleAspIleLysMetAlaIleProGluLeuMetArgGluAspV 484
 1381 AGGCATTGATCTATCAACCTTATCATACAGATCTCATGCGGGAAGATG 1430
 484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
 1431 TTCAATTGTTCATGCTTGGATCTGGTCACCCAGAGCTTGAAGATTGGATG 1480
 501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
 1481 AGATCTACAGAGTCGATCTTCAGAGATAAATTCCTGGGATGGGTGGATT 1530
 517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
 1531 TAGTGTTCAGTTTCCCAACGAATAACTGCGGCTCCGATATATTGTTAA 1580
 534 eProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
 1581 TGCCATCCAGATTGCAACCTTGTGGTCTCAATCAGCATATATGCTATGCAG 1630
 551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
 1631 TATGGCAGATTCTGTGTGTCATGCAACTGGGGCCTTAGAGATACCCT 1680
 567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
 1681 GGAGAACTTCAACCTTTCGGTGAGATGGAGAGAGGGTACAGGGTGG 1730
 584 lApeSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
 1731 CATTCGACCCCTTAACCACAGAAACATGTTCTGGACATTCGGAACCTGCA 1780
 601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
 1781 ATATCTACATACAGGAACAACAAGTCTCTCTCTGGGAAGGCCTAATGAAGCG 1830
 617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSer 633
 1831 AGGCATGTCAAGAGACTTCACGTGGGACCATGCGCGCTGAACCATACGAA 1880
 634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG 650

Fri Mar 29 09:26:44 2002

us-09-674-824-2.rnmp

1881 AATCTTCAGTGGCCCTTCATCGACCCATGTCATGTAAATAA... 1928

650 yLeuGlyArgSerLysCysGluSerPro 659

1929AGGACCAAGTCGGTTCCT 1949

seq_name: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:US-09-743-980-12

seq_documentation_block:

; Sequence 12, Application US/09743980
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147-A
; CURRENT APPLICATION NUMBER: US/09/743,980
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 060/094,436
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Zea mays
US-09-743-980-12

alignment_scores:
Quality: 2806.50 Length: 660
Ratio: 4.717 Gaps: 6
Percent Similarity: 90.152 Percent Identity: 80.909

alignment_block:
US-09-674-824-2 x US-09-743-980-12 ..

Align seg 1/1 to: US-09-743-980-12 from: 1 to: 2008

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuAlaAs 20
14 TCGCCGCTGGCGCGCGCTGCCCTC..... 40
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
41 .CTCGCGCGCGCGCTGGCGCGCGCTGCCCGCGCGCGCGCGCG 89
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
90 GCGCGCTCCAGCGGTGCTGGCGCGCGCGCTGCCCGCGCGCGCGCG 139
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
140 GAGGGCGCGCGCGCGCGCGCGCTGCCCGCGCGCTGCCCGCGCGCT 189
67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerPro 84
190 CGTGGCGCGCTTCCTCGCG...CGCGCGCGCGCGCGCGCGCGCG 236
84 laProThrGlnProProProProAlaGlyValGlyGluLeuAlaPro 100
237 CATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 283
101 AspLeuLeuGluGlyLeuAlaGluAspSerLeuAspSerLeuLeuLe 117
284 ...CTGACCACTGAAGGATTCCTCAAGGTTCCATCGATAACACAGTAGT 330
117 lAlaAlaSerGluGlnAspSerGluLeuMetAspAlaAsnGluGlnPro 134
331 TGTGCAAGTACAGAGATCTGAGATTCTGTTGGAAGGACGAGCTC 380
134 lAlaLysValThrArgSerLeuValPheValThrGlyGluAlaPro 150

381 GAGCTAAAGTAACACAAAGCATTTGTTTGTAAACGGGGAAGCTTCCT 430
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
431 TATGCAAGTCTGGGGTCTAGGAGATGTTGTGGTTTCATTGCCAGTTGC 480
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
481 TCTTGCTGCTGCTGCTACCGTGTGATGTTGTAATGCCAGATATTAA 530
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
531 ATGGTACCTCCGATAGAATATGCAATATGCAATATGCAATATGCAAT 580
201 lLeLysileProCysPheGlyGlySerHisGluValThrPhePheHisG 217
581 ATTCCGATTCATGCTTTGGCGGTGAACATGAAGTTACCTTCTCCATGA 630
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
631 GTATAGAGATTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
681 GACTGGAAATTTATATGGAGATAAGTTGGTCTTTGGTGTATAATCAG 730
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIle 267
731 TTCAGATACACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
267 uGluLeuGlyGlyTyrTyrTyrGlyGlnAsnCysMetPheValValAsnA 284
781 TGAATGGGAGGATATATTTATGGACAGATTCATGCTTTGTTGTTCA 830
284 sPTpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
831 ATTGCATGCCAGTCTAGTCCAGTCTCTTCTGCTGCAAAATATAGACCA 880
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
881 TATGGTGTATAAAGACTCCCGCAGCATCTTCTGCTGCAAAATATAGAC 930
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 334
931 ACATCAGGCTGACAGCTGCAAGCACATATCTGACCTTGGGTGGCCAC 980
334 roGluTyrTyrGlyAlaLeuGluTrpValPheProGluTyrAlaArgArg 350
981 CTGAATGGTATGGAGCTCTGGAGTGGGTATTTCCCTGAATGGCGAGGAG 1030
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaVal 367
1031 CATGCCCTTGACAGGCTGAGCAGTTAAATTTTGAAGGTCAGCTGT 1080
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluVal 384
1081 GACAGCAGATCGAATCGTACTGTGAGTGGGTATTTTCGTGGGAGGTCA 1130
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1131 CAACCTGCTGAAGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAA 1180
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsn 417
1181 AGTGTATTAAACGGAATTTGTAATGGAATTTGACATTAATGATTTGA 1230
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspLeuSerG 434
1231 TGCACAGACAAATGATCCCTGCTGCTATTTCTGTTGATGACCTCTCTG 1280
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1281 GAAAGCCAAATGTAAGGTCATTCAGAGGAGCTGGGTTTACCTATA 1330

```
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1331 AGGCTGATGTTCTCTGATGGCTTATTGGAAGGTTGATTTATCAGAA 1380
467 sGlyIleAspLeuIleGlyMetAlaIleProGluLeuMetArgGluAspY 484
1381 AGGCATTGATCTCATCACTATCATACCATCATCATCATCATCATCAT 1430
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
1431 TTCAATTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATG 1480
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
1481 AGATCTACAGATCGATCTTCAAGGATAAATTCGTGGATGGTGGATT 1530
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1531 TAGTGTTCAGATTTCACCGAATAACTCCGGCTGCGATATATTGTTAA 1580
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1581 TGCCATCCAGATTGCAACCTTGTGTCTCAATCAGCTATATCTATGCGAG 1630
551 TyrGlyThrValProValHisGlyThrGlyLeuArgAspThrVa 567
1631 TAGTGCACAGTTCTCTGTCTCCATCAACTGGGGGCTTAGAGATACCGT 1680
567 lGluThrPheAsnProPheGlyAlaGlyGlyGluGlyThrGlyTrpA 584
1681 GGAGAACTTCAACCTTCCGTGAGAAATGTTGGACATTCGAACTGCA 1730
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1731 CATTCGCACCCCTAACACACAGAAACATGTTGGACATTCGAACTGCA 1780
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1781 ATATCTACATACGAGGAACAAAGTCCTCCCTGGGAAGGCTTAATGAAGC 1830
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
1831 AGGCATGTCAAAGACTTCACGTGGGACCATGCCGCTGAACAATACGAAC 1880
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGl 650
1881 AAATCTCCAGTGGGCTTCATCGATCGACCTATGTCATGTAATAAAAA.. 1928
650 yLeuGlyArgSerLysCysGluSerPro 659
1929 .....AGGACCAAGTGGTGGTTCCT 1949
```

```
seq_name: /cgn2_5/ptodata/1/pna/US6009_COMB.seq:US-60-094-436-12
```

```
seq_documentation_block:
```

```
; Sequence 12, Application
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Modification of Starch Biosynthetic Enzyme Gene
; FILE REFERENCE: BB-1147-p1
; CURRENT APPLICATION NUMBER: US/60/094,436
; CURRENT FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 12
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Zea mays
US-60-094-436-12
```

```
alignment_scores:
  Quality: 2806.50      Length: 660
  Ratio: 4.717         Gaps: 6
  Percent Similarity: 90.152  Percent Identity: 80.909
alignment_block:
US-09-674-824-2 x US-60-094-436-12 ..
Align seg 1/1 to: US-60-094-436-12 from: 1 to: 2008
4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuAspAlaAs 20
14 TCGGCGGTGGCGCGCGCTGCTCCTC..... 40
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
41 .CTCGCGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCG 89
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
90 GCGCGCTCCAGCGGTGCTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCG 139
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
140 GAGGGGCGCGCGCGCGCTGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCT 189
67 uValProGlyPheLeuAlaProProProProProAlaProAlaGlnSerProA 84
190 CGTGCGCGGTCTCTCGCG...CGCGCGCGCGCGCGCGCGCGCGCGCG 236
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
237 CATCGACGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCTGCGCGCG 283
101 AspLeuLeuLeuGlyIleAlaGluAspSerIleAspSerIleIleVa 117
284 ...CTCGAAGCTGAAGGATGCTGAAGGTTCCATCATCAACACAGTAGT 330
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
331 TGTGGCAAGTGACCAAGATTCAGATGTTGGTTGGAAAGGACCAAGCTC 380
134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
381 GAGCTAAAGTAACAAAGCATTTGCTTTGTAACCGCGGAAAGCTTCCTC 430
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
431 TATGCAAGCTCGGGGCTTAGAGATGTTTGTGTTTTCATTGCCAGTTGC 480
167 alLeuAlaAlaArgGlyHisArgValMetValMetProArgTyrLeuA 184
481 TCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
531 ATGGTACCTCCCATAGAATATGCAAAATGCAATTTTACACAGAAACAC 580
201 lLeLysIleProCysPheGlySerHisGluValThrPhePheHisGl 217
581 ATTCGATTCCATGCTTTTGGCGGTGAACATGAAGTTACCTTCTTCATGA 630
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
631 GTATAGAGATTACAGTTGACTGGTGTGTTTGTGATCATCCCTCATATACA 680
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
681 GACCTGGAATTTATATGAGATAAGTTTGGTCTTTTGGTGAATATCAG 730
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
```



```
||||: :||| ||| |||:|||||
101 ACCCCACGGGTGAGCGGCATCGACGCCGCCCGTGCCTCCGACCGCGC 150
95 ValGlyGluLeuAlaProAspLeuLeuGluGlyIleAlaGluAspSe 111
:||||:||||: ||| |||:|||||
151 CTGGGGGACCTCGGT.....CTCGAACCTGAAGGGATTGCTGAAGGTC 194
111 rIleAspSerIleValAlaAlaSerGluGlnAspSerGluIleMeta 128
|||||:||||: :||||:||||: ||| |||:|||||
195 CATCGATAACACAGTAGTTGTGGCAAGTGAGCAAGATTCTGAGATTGG 244
128 sAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPheVal 144
:||||:||||: :||||:||||: ||| |||:|||||
245 TTGGAAGAGCAGCTCGACCTAAAGTAACACAAGCATTGCTCTTTGTA 294
145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCy 161
|||||:||||: :||||:||||: ||| |||:|||||
295 ACCGCGGAAGCTTCCTCTATGCAAGTCTGGGGCTAGGAGATGTTG 344
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValV 178
|||||:||||: :||||:||||: ||| |||:|||||
345 TGGTTTCATTGCGCAGTTGCTTGTGCTGCTGCTGGTCAACGCTGATGG 394
178 alMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAla 194
|||||:||||: :||||:||||: ||| |||:|||||
395 TAATGCCAGATATTAAATGGTACCTCCGATTAAGAATTATGCAATGCA 444
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisG1 211
|||||:||||: :||||:||||: ||| |||:|||||
445 TTTTACACAGAAAAACACATTTCGGATCCATGCTTTGGCGGTGACATGA 494
211 uValThrPhePheHisGluTyrArgAspAsnValAspTrpValPheValA 228
|||||:||||: :||||:||||: ||| |||:|||||
495 AGTTACCTCTTCATGATAGTATAGATTCAGTTGACTGGGTGTTGTTG 544
228 sPHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGly 244
|||||:||||: :||||:||||: ||| |||:|||||
545 ATCATCCTCATATCACAGACCTCGAATTTATATGGAGATAAGTTGGT 594
245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystyrAlaAlaCy 261
|||||:||||: :||||:||||: ||| |||:|||||
595 GCTTTTGGTGATATACAGTTCAGATACACACTCCTTTGCTATGCTGATG 644
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnC 278
|||||:||||: :||||:||||: ||| |||:|||||
645 TGAGGCTCCCTTGTATCTCGAATGGGAGGATATATTTATGGACAGATT 694
278 yMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294
|||||:||||: :||||:||||: ||| |||:|||||
695 GCATGTTTGTTCATGATTGGCATGCCAGTCTAGTGCCAGCTTCCTTCT 744
295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311
|||||:||||: :||||:||||: ||| |||:|||||
745 GCTGCAAAATATAGACCATATGCTGTTTATAAAGACTCCCGCAGCATCT 794
311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328
|||||:||||: :||||:||||: ||| |||:|||||
795 TGTAAATACATATTAGCATCAGGCTGAGAGCCCTGCAAGCACATATC 844
328 roAspLeuGlyLeuProProGluTrpTyrGlyAlaLeuGluTrpValPhe 344
|||||:||||: :||||:||||: ||| |||:|||||
845 CTGACCTTGGGTTGCCACCTGAATGATGAGCTCTGGAGTGGGTATTC 894
345 ProGluTrpAlaArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361
|||||:||||: :||||:||||: ||| |||:|||||
895 CCTGAATGGCGAGGAGGCATGCTTGTACAAAGGCTGAGCAGTAAATTT 944
361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnC 378
|||||:||||: :||||:||||: ||| |||:|||||
945 TTTGAAAGTGCGAGTTGTGACACACATCGAATCGTCACTGTCAGTAAGG 994
378 lyTyrSerTrpGluValThrAlaGluGlyGlnGlyLeuAsnGlu 394
|||||:||||: :||||:||||: ||| |||:|||||
```

```
995 GTTATTCTGGGAGCTCACACTGCTGAAGTGGACAGGGCTCAATGAG 1044
395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411
|||||:||||: :||||:||||: ||| |||:|||||
1045 CTCCTTAAGCTCCAGAAAGAGTGTATTAACGGAATTGTAATGAATTGA 1094
411 pIleAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428
|||||:||||: :||||:||||: ||| |||:|||||
1095 CATTAAATGATTGGAAACCTGCCACAGACAAATGTATCCCTGTCTATT 1144
428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
|||||:||||: :||||:||||: ||| |||:|||||
1145 CTGTTGATGACCTCTCTGGAAGGCCAATGTAAAGGTGCTATTCAGAG 1194
445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG1 461
|||||:||||: :||||:||||: ||| |||:|||||
1195 GAGCTGGGTTTACCTATAAGGCCCTGATGCTCTCTGATTGGCTTATTG 1244
461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478
|||||:||||: :||||:||||: ||| |||:|||||
1245 AAGTTGGATTATCAGAAAGCATTTGATCTCATTTCAACTTATCATACC 1294
478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494
|||||:||||: :||||:||||: ||| |||:|||||
1295 ATCTCATGCGGGAAGATGTTCAATTTGTCATGCTTGGATCTGGTGACCC 1344
495 IlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspLysPh 511
|||||:||||: :||||:||||: ||| |||:|||||
1345 GAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAAT 1394
511 eaArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAlaG 528
|||||:||||: :||||:||||: ||| |||:|||||
1395 TCGTGGATGGTGGATTTAGTGTCCAGTTTCCCACCCGAATAACTGCCG 1444
528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544
|||||:||||: :||||:||||: ||| |||:|||||
1445 GCTCGCATATATTGTTAATGCCATCCAGATTCGAAACCTTGTGGTCTCA 1494
545 GluLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrG1 561
|||||:||||: :||||:||||: ||| |||:|||||
1495 CAGCTATATGCTATGCAAGTATGGCACAGTTCCTGTTCTCATGCCAATGG 1544
561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578
|||||:||||: :||||:||||: ||| |||:|||||
1545 GGGCTTAGAGATACCGTGGAGAACTTCAACCTTTCGCTGAGAATGGAG 1594
578 luGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLe 594
|||||:||||: :||||:||||: ||| |||:|||||
1595 AGCAGGTCACAGGTTGGGCTTCGCCACCCCTAACCCACAGAAAACATGTT 1644
594 uTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611
|||||:||||: :||||:||||: ||| |||:|||||
1645 GTGCAATTCGCAACTGCAATATCTACATACAGGGAACACAAGTCTCTCT 1694
611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627
|||||:||||: :||||:||||: ||| |||:|||||
1695 GGAAGGGCTAATGAAGCGAGGCAATGTCAAAACACTTCACGTGGGACCAT 1744
628 Ala 628
|||
1745 GCC 1747
```

```
seq_name: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:us-09-625-406-12
seq_documentation_block:
; Sequence 12, Application US/09625406
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
```

```

; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/625,406
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/941,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1752 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1752
; US-09-625-406-12

```

```

alignment_scores:
  Quality: 2713.50      Length: 585
  Ratio: 4.916          Gaps: 3
  Percent Similarity: 94.359  Percent Identity: 86.325

```

```

alignment_block:
US-09-674-824-2 x US-09-625-406-12 ..
Align seg 1/1 to: US-09-625-406-12 from: 1 to: 1752
45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnG1 61
|||||
4 GTCGCGAGCTGACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 53
61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78
|||||
54 GCTGCTGGCGCGCGCGCTGCTGCGCGCGCTGCTGCGCG...CCGCCGCG 100
78 laProAlaGlnSerProAlaProThrGlnProProLeuProAlaGly 94
|||||
101 AGCCACAGCGGTGAGCGCGCATCGACGCCCGCGCGCGCGCGCGCGCG 150
95 valGlyGluLeuAlaProAspLeuLeuGluGlyIleAlaGluAspSe 111
|||||
151 CTGGGGGACCTCGT.....CTCGAACCTGAAGGATTCGTCGAAGGTT 194
111 rleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMe 128
|||||
195 CATCGATAACACAGTAGTGTGGCAAGTGCAGCAAGATTCGAGATTGT 244
128 spAlaAsnGluProGlnProGlnAlaLysValThrArgSerIleVal 144
|||||
245 TTGGAAGGAGCAAGCTCGAGCTAAAGTACACAAAGCATTCGCTTTGTA 294

```

```

145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValcy 161
|||||
295 ACCGCGAAGCTTCTCTTATGCAAAAGCTGGGGTCTAGGAGATGTTG 344
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetVal 178
|||||
345 TGGTTCATTCGCAAGTCTCTTGTGCTGCTGGTCAACGTTGATGTTG 394
178 alMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAla 194
|||||
395 TAAATGCCCCAGATATTAAATGGTACCTCCGATAAGAATTTATGCAAAATGCA 444
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisG1 211
|||||
445 TTTTACACAGAAAACACATTCGGATTCCATGCTTTGGCGGTGAACATGA 494
211 uValThrPhePheHisGluTyrArgAspAsnValAspTrpValPheVala 228
|||||
495 AGTTACCTTCTTCATGAGTATAGATTCAGTTGACTGGGTGTTTGTG 544
228 spHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGly 244
|||||
545 ATCATCCCTCATATCACAGACCTGGAATTTATATGGAGATAAGTTGGT 594
245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCy 261
|||||
595 GCTTTTGGTGATAATCAGTTTCAGATACACACTCTTCTGCTATGCTGCATG 644
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnC 278
|||||
645 TGAGGCTCTCTTGTGATCCTTGAATGGGAGGATATATTTATGGACAGAA 694
278 ysMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294
|||||
695 GCATGTTTGTGCAATGATTGGCATGCCAGTCTAGTCCAGCTCCTTCT 744
295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311
|||||
745 GCTGCAAAATATAGACCATATGTTTATAAAGACTCCCGCAGCATCTCT 794
311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328
|||||
795 TGTATACATATTTAGCACATCAGGCTGTAGAGCTGCAAGCACATATC 844
328 roAspLeuGlyLeuProProGluTrpTyrGlyAlaLeuGluTrpValPhe 344
|||||
845 CTGACCTTGGGTTGCCACCTGAATGGTATGGAGCTCTGGAGTGGGTATTC 894
345 ProGluTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361
|||||
895 CCTGAATGGCGAGGAGGATCCCTTGCACAGGCTGAGGCAGTAAATTT 944
361 eLeuLysGlyAlaValAlaValThrAlaAspArgIleValThrValSerGlnG 378
|||||
945 TTTGAAAGGTGCAGTTGTGACAGCAGATCGAATCGTACTGTGCTAGTAA 994
378 lyTyrSerTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394
|||||
995 GTTATTCTGGGAGGTTCACAACTGCTGAAGGTGGACAGGCGCTCAATGAG 1044
395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411
|||||
1045 CTCCTTAAGCTCCAGAAAGAGTGTATTAACGGAATTTGTAATGGAATGA 1094
411 pileAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428
|||||
1095 CATTAATGATTGGAACCTGCGCACAGACACAAATGATCCCTCTGCTATT 1144
428 ervAlaAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
|||||
1145 CTGTTGATGACCTCTCTGGAAGGCGCAATGTAAGAGGTGCATTTGCAAG 1194

```

```
445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleGl 461
|||||
1195 GAGCTGGGTTTACCTATAGGCGCTGATGTTCTCTGATGGCTTTATGG 1244
|||||
461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478
|||||
1245 AAGGTTGGATTATCAGAAGGCGATTGATCTCATTTCAACTTATCATACCAG 1294
|||||
478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494
|||||
1295 ATCTCATCGGGAAGATGTTCAATTTGTGTCATGCTTGGATCTGGTGAACCA 1344
|||||
495 llePheGluGlyTyrMetArgSerThrGluSerSerTyrLysAspLysPh 511
|||||
1345 GAGCTTGAAGATTGGATGAGATCTACAGATCGATCTTCAAGGATAAATT 1394
|||||
511 eArgGlyTyrValGlyPheSerValProValSerHisArgIleThrAlaG 528
|||||
1395 TCGTGGATGGGTTGGATTAGTTAGTTTCCAGTTTCCACCGAATAACTGCCG 1444
|||||
528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544
|||||
1445 GCTGGATATATGTTATGATGCATCCAGATTCGAACCTTGGTCTCAAT 1494
|||||
545 GlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrGl 561
|||||
1495 CAGCTATATGTCATGTCAGTATGGCAGATTCCTGTTGTCCATGCAACTGG 1544
|||||
561 yGlyLeuArgAspThrValGluThrPheAsnPropheGlyAlaLysGlyG 578
|||||
1545 GGGCTTAGACATACCGTGGAGAACTTCAACCCCTTTCGGTGAGAAATGGAG 1594
|||||
578 luGluGlyThrGlyTyrAlaPheSerProLeuThrValAspLysMet.Le 594
|||||
1595 AGCAGGTACAGGCTGGGCAATTCACCCCTTAACACAGAAACATGTTT 1644
|||||
594 uTyrAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611
|||||
1645 GTGGACATGGCAACTGCAATATCTACATACAGGGAACACAAAGTCCCT 1694
|||||
611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTyrAspHis 627
|||||
1695 GGGAAAGGCTAATCAAGCGAGCATGTCAAAGAACTTCACGTGGGACCAT 1744
|||||
628 Ala 628
|||
1745 GCC 1747
```

seq_name: /cgn2_6/ptodata/1/pna/US090_COMB.seq:US-09-077-564-1

seq_documentation_block:

Sequence 1, Application US/09077564

GENERAL INFORMATION:

APPLICANT: Knight, Mary E.

APPLICANT: Keeling, Peter L.

TITLE OF INVENTION: Modification of Starch Synthesis in

TITLE OF INVENTION: Plants

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: ZENECA Ag Products

STREET: 1800 Concord Pike

CITY: Wilmington

STATE: DE

COUNTRY: USA

ZIP: 19850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/077,564

FILING DATE: 14-DEC-1998

```
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/02990
FILING DATE: 04-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524938.9
FILING DATE: 06-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hohenschutz, Liza D.
REGISTRATION NUMBER: 33,712
REFERENCE/DOCKET NUMBER: SEE 45052/UST
TELEPHONE: (302) 886-1699
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2992 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: NUMBER 1
US-09-077-564-1

alignment_scores:
Quality: 2655.50 Length: 745
Ratio: 4.304 Gaps: 11
Percent Similarity: 82.819 Percent Identity: 71.544

alignment_block:
US-09-674-824-2 x US-09-077-564-1 ..
Align seg 1/1 to: US-09-077-564-1 from: 1 to: 2992

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
|||||
614 TCGCGCGTGGCGCGCGGTGCTC.....CTCCTCGCGCG 648

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAlaLeu 37
|||||
649 GNGCGCTGGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 698

37 rArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPro 53
|||||
699 AGCGGTGCTGCGCGCGCGGTGCTCGCGAGCTGACGAGGAGGCGCC 748

54 AlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLeuValProGl 70
|||||
749 GCSCCGCGCGCGCTGCCACCGCGCTGCTGGCGCGCGCGCGCGCG 798

70 yPheLeuAlaProProProProAlaProAlaGlnSerProAlaProThrG 87
|||||
799 CTTCCTCGCG...CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 845

87 lnProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeu 103
|||||
846 CGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889

104 LeuGluGlyIleAlaGluAspSerIleAspSerIleValAlaAla 120
|||||
890 CCTGAAGGATTGCTGAAGGTTCCATCGATACACAGTAGTTGTGGCAAG 939

120 rGluGlnAspSerGluIleMetAspAlaAsnGlnProGlnAlaLysV 137
|||||
940 TGAGCAAGATTCTGAGATTGTGGTGAAGGAGGAGCAAGCTCGAGCTAAG 989

137 alThrArgSerIleValPheValThrGlyGluAlaAlaPro.TyrAla 153
|||||
990 TAACACAAAGCATTTGCTTTTAAACCGCGGAGGCTTCTCTTAATCGAA 1039

153 sSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAla 170
|||||
```

[illegible]

470 spLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPhe 486
 1990 ATCTCAATCAACTTATCATACCAAGATCTCATGCGGAAGAATGTTCAAATTT
 487 ValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerTh 503
 2040 GTCAATGCTTGGATCTGGTGACCCAGACCTTGAAGATTTGGATGAGATCTAC 2089
 503 tGluSerSerTrpLysAspLysPheArgGlyTrpValGlyPheSerValP 520
 2090 AGAGTCGATCTTCAAGGATAAATTTCTGGATGGGTGGATTTAGTGTTC 2139
 520 tGValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSer 536
 2140 CAGTTTCCACCGAATAACTCGCGCTCGCATATATTTGTTAATGCCATCC 2189
 537 ArgPheGluProCysGlyLeuAsnGlnLeuTrpAlaMetGlnTrpGlyTh 553
 2190 AGATTTCGAACCTTTGTGTCTCAATCAGCTATATGCTATGCAGTATGGCAC 2239
 553 rValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrP 570
 2240 AGTTCCTGTGTGCCATGCACTACTGGGGCCTTAGAGTACCGTGGCAACT 2289
 570 heAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSer 586
 2290 TCAACCCCTTTCGGTTCAGAAATCGAGAGCAGGGTACAGCGGTGGCGATTTCGCA 2339
 587 ProLeuThrValAspLysMet_LeuTrpAlaLeuArgThrAlaMetSerT 603
 2340 CCCCATACCACAGAAACATGTTGTGGACATTCGCAACTGCAATATCTA 2389
 603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
 2390 CATACAGGAACACACAAGTAATAATGGGAAGGCCTAATGAAGCCAGCCAGCATG 2439
 620 ThrLysAspHisThrTrpAspHisAlaPro_SerSerThrSerArgSerS 636
 2440 TCAAAAGAGTTCACGTGGGACCATGCCCTGACCAATACGAACAAATCTT 2489
 636 exSerGlyProSerTrpThrAsnProThrSerCysArg_ArgGlyLeuGlu 652
 2490 CCAGTGGGCTTCATCGGATGCACCC...GATGTTCAATGGAAAAAGGG 2536
 652 yArgSerLysCysGlySerProSerAlaLeuLysThrSerSerSerP 669
 2537 ACCAAAGTTGGTGGTTCTTGAAGATTTACAGTTTCATCTCATATAGTA 2586
 669 heArgGlyProGluGlyTrpPro...CysThrLeuArgCysProAla 683
 2587 AGCTGAATGATGAAGAAACCCCTGTACATTACATGAAGGCAGACCGG 2636
 684 ThrValGlnSerGlnCysAlaCysLeuLeuTrpPhe..... 695
 2637 CTATTGGCTCCATGCTCCAATGTCGTGTTGGCTTCGCTTCGCTCGATGG 2686
 695 695
 2687 ACCGGATGCAGTGAGGAATCCAGNCGAAGCAGAGTTTGAAGGATAGGAA 2736
 696AlaGlySerArgThrTrpAspGlyCysAlaAlaAlaValThr 710
 2737 GGGGAGCTGAAGC...AGTCAGCGCAGCAGCGCAAGCCCTTCGCCGTAAAT 2783
 711 AlaSerGlyGlyArgGlnLeuGlnPhe 719
 2784 TCATATGGAAACAAGCTGGAGTCAGTTT 2810

```
seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq: US-09-654-617-4517533
seq_documentation_block:
: Sequence 451753, Application US/09654617
```


; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
 ; TITLE OF INVENTION: Annotated Plant Genes
 ; FILE REFERENCE: 38-21(15097)D
 ; CURRENT APPLICATION NUMBER: US/09/654,617
 ; CURRENT FILING DATE: 2000-09-05
 ; NUMBER OF SEQ ID NOS: 463173
 ; SEQ ID NO 451753
 ; LENGTH: 2216
 ; TYPE: DNA
 ; ORGANISM: Sorghum bicolor
 US-09-654-617-451753

alignment_scores: Quality: 2578.00 Length: 628
 Ratio: 4.628 Gaps: 9
 Percent Similarity: 88.694 Percent Identity: 78.662

alignment_block:
 US-09-674-824-2 x US-09-654-617-451753 ..
 Align seg 1/1 to: US-09-654-617-451753 from: 1 to: 2216

107 lleAlaGluAspSerIleAspSerIlelleValAlaAlaSerGluGlnAs 123
 1 ATTGCTGAAGGTTCCATGCATGAGACAGTAGTTGTGGCAAGTGAGCAGA 50
 123 pSerGluIleMetaspAlaAsnGluGlnProGlnAlaAlaLysValThrArgS 140
 51 TCTGAGATCGTGGTTGGAAGGAGCAGCTCGAGCTAAAGTAAACACAAA 100
 140 erIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGly 156
 101 GCATTGTCTTTGTAACCTGGCGAAGCTTCTCCTTATGCAAGCTCGGGGT 150
 157 LeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyH1 173
 151 CTAGAGAGATGTTGGTGTTCATGCCAGTGTGCTTCTGCTCGTGCTCA 200
 173 sArgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysA 190
 201 CCGTGTGATGTTGTAATGCCACATATTTAAATGGTACCTCTGTATAGA 250
 190 snTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPhe 206
 251 ATTACGCAAAATGCAATTTACACAGAAAGCAGACATTCCGATTCATGCTT 300
 207 GlyGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAs 223
 301 GCGGTGAACATGAAGTTACCTTTTCCATGAGTACAGAGATTCAAGTTGA 350
 223 pTrpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrG 240
 351 CTGGGTGTTTGTGATCATCCCTCATATCACAGACCTGGAAATTTATATG 400
 240 lyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeu 256
 401 GAGATAAGTTTGGTCTTTGGTGAATACAGTTTCAGATACAGCGCTCCT 450
 257 CysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyTyrIle 273
 451 TGCTATGCTGCATGTGAGGCTCTTTGGTCTTGAATTTGGGAGGATATAT 500
 273 eTyrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuV 290
 501 TTTATGACAGAATTTGATGTTTGTGTAATGATTGCCATGCCAGCTAG 550
 290 alProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAsp 306
 551 TGCCAGCTCTTCTGCTGCAAAATATAGACCATATGTTGTTTATAAAGAC 600

307 SerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluPr 323
 601 TCCCGCAGCATCTCTGTAAATACATAATTTAGCACATCAGGGGTAGAGCC 650
 323 oAlaSerThrTyrProAspLeuGlyLeuProProGluTyrPyrGlyAlaL 340
 651 TGCAAGCACATATCTGACCTGGCTTGGCTGCCACCTGAATGGTATGAGCTC 700
 340 euGluTyrValPheProGluTrpAlaArgArgHisAlaLeuAspLysGly 356
 701 TGGAGTGGGTATTCCTGAATGGCAAGGAGGCATGCCCTTGACAAGGGT 750
 357 GluAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleVa 373
 751 GAGCAGTTAATTTTGAAGGTGCAGTTGTGACACAGCATCGAATGT 800
 373 lThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGly 390
 801 GACTGTGAGTAAGGGTTATTCTGGGAGGTCACAACCTGCTGAAGGTGGAC 850
 390 lnGlyLeuAsnGluLeuSerSerArgLysSerValLeuAsnGlyIle 406
 851 AGGCTCTCAATGAGCTCTTAAGCTCCCGAAGAGTGTATTAAACGGAATT 900
 407 ValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLe 423
 901 GTAATGGAATTGACATTAATGATTGGAACCTTGCAGCGGACAAATGTAT 950
 423 uProHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysA 440
 951 CCCTTGTCTATTATCTGTGATGACCTCTCTGGAAGGCCAAATGTAAAA 1000
 440 laGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeu 456
 1001 GTCATATGCAGAAAGAGCTGGGTTTACCTATAAGGCTGAAGTTCCTCTG 1050
 457 IleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleY 473
 1051 ATTGGTTTTATTGGAAGATTGGATTATCAGAAAGGCATTGATCTCATTC 1100
 473 sMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuG 490
 1101 ACTTATCATACCATCTCATCGGGAGCAGCTTCAATTTGTCATGCTTG 1150
 490 lySerGlyAspProIlePheGluGlyTyrMetArgSerThrGluSerSer 506
 1151 GATCTGGTACCCAGAGCTCGAAGACTGGATGAGATCTACAGAGTCGGAC 1200
 507 TyrLysAspLysPheArgGlyTyrValGlyPheSerValProValSerHi 523
 1201 TTCAAGGATAAATTTCTGTGGATGGGTTGGATTAGTTTCCAGTTTCCCA 1250
 523 sArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluP 540
 1251 CCGAATAACTGCGGCTGCGATATATTGTTAATGTCATCCAGATTCGAAC 1300
 540 roCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVal 556
 1301 CTTGTGGTCTCAATCAGCTATATGATGTCAGTATGGCAGACAGTCTCTGT 1350
 557 ValHisGlyThrGlyLeuArgAspThrValGluThrPheAsnProPh 573
 1351 GTCCATGCAACTGGGGGCTTAGAGATACTGTGGAGAACTTCAACCCCTT 1400
 573 eGlyAlaLysGlyGluGlyThrGlyTyrAlaPheSerProLeuThrV 590
 1401 CGGTGAGAAATGGAGAACAGGGTACAGGTTGGGCACTTCGACCCCTTAA 1450
 590 alAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgG 606
 1451 CGGAAACATGTTCTGTGACATTCGGAACCTGCAATTTCCACATACAGGA 1500
 606 uHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspH 623

```
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1501 GCACAAATCTCTCGGAGGGCTCATGAGAGGGGATGTCACAAAGACT 1550
623 isThrTrpAspHisAlaPro..SerSerThrSerArgSerSerGlyPr 639
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1551 TCACGTGGGACCATGCCGCTGACAAATACGAACAAATCTTCCAGTGGGC 1600
639 oSerTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLysC 656
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1601 TTCATCGATCGACCCCTATGTCATGATAAGGGGAC..... 1634
656 ysGlu.SerProSerAlaLeuLysThr...SerSerSerSerPheArgG 671
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1635 ..CACAAGTGGTGTCTCTTAAAGATCTTCAGTTCTTCATCCTACATGG 1682
671 yProGlu.GlyTyProCysThrLeuArgCysProAlaThrValGluSer 687
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1683 AAGGATGAGAAACCCCTGTATACATTACAT..... 1711
688 GlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTrpAspGlyCy 704
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1712 .....GGAGCGGACC.....GGCTG 1728
704 sAlaAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnPheTrpG 721
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1729 TTGGGTGGCTCCAAATGTCAGGTTAGCCTTACGGAGCCTGCTGCTTGGC 1778
721 yLeuArgLysGlyCysAlaAlaGly 729
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1779 TGGTT.....TGTCACCAAGGA 1795
```

seq_name: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:US-09-684-016-451753

```
seq_documentation_block:
; Sequence 451753, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 451753
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Sorghum bicolor
US-09-684-016-451753
```

```
alignment_scores:
  Quality: 2578.00      Length: 628
  Ratio: 4.628          Caps: 9
  Percent Similarity: 88.694  Percent Identity: 78.662
```

alignment_block:
US-09-674-824-2 x US-09-684-016-451753 ..

Align seg 1/1 to: US-09-684-016-451753 from: 1 to: 2216

```
107 IleAlaGluAspSerIleAspSerIleValAlaAlaSerGluClnAs 123
1 ATTGCTGAAGTTCATCGATGACAGAGTAGTTGTGGCAAGTGAGCAAGA 50
123 pSerGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgS 140
51 TTCTGAGATCGTGTGGGAAGGAGCAAGCTCCAGCTAAGTAAGTAACACAA 100
140 erIleValPheValThrGlyGluAlaAlaProTyrrAlaLysSerGlyGly 156
101 GCATTGCTTTGTAACCTGGCGCAAGCTTCCTTATGCAAAAGTCTGGGGGT 150
```

```
157 LeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHI 173
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
151 CTAGGAGATGTTTGGTTCAITGGCAGTTGCTTCTTGGTCTCGTGGTCA 200
173 sArgValMetValValMetProArgTyrrLeuAsnGlySerSerAspLysA 190
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
201 CGCTGTGATGGTGTAAATGCCAGATATTTAAATGGTACCTCTGATAAGA 250
190 sTyrrAlaLysAlaLeuTyrrThrAlaLysHisIleLysIleProCysPhe 206
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
251 ATTACGCAATGCAATTTACAGAAAACACATTCGGATTCCATGCTTT 300
207 GlyGlySerHisGluValThrPhePheHisGluTyrrArgAspAsnValAs 223
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
301 GGCGGTGAACATGAAGTTACCTTTTCCATGAGTACAGAGATTCACTTGA 350
223 pTrpValPheValAspHisProSerTyrrHisArgProGlySerLeuTyrrC 240
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
351 CTGGGTGTTTGTGATCATCCCTCATACAGACCTGGAAATTTATATG 400
240 yAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrrThrLeuLeu 256
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
401 GAGATAAGTTTGGTCTTTGGTGATAATCAGTTTACAGATACACGCTCCTT 450
257 CysTyrrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyClyTyrrI 273
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
451 TGCTATGCTGCATGTGAGGCTCCTTTGGTCTTGAATTTGGGAGGATATAT 500
273 eTyrrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuV 290
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
501 TTATGGACAGAAATTCATGTTTGTGTGAATGATTGGCATGCCAGCTTAG 550
290 alProValLeuLeuAlaAlaLysTyrrArgProTyrrGlyValTyrrArgAsp 306
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
551 TGCAGTCCCTTCTTGTGCAAAATATAGACCATATGTTTATAAGAC 600
307 SerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluPr 323
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
601 TCCGCGCAGCATTCCTGTAATACATAATTTAGCACATCAGGGTGTAGAGCC 650
323 oAlaSerThrTyrrProAspLeuGlyLeuProProGluTyrrTyrrGlyAlaL 340
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
651 TGCANGCACATATCCTGACCTTGGGTTGCCACCTGAAATGTATGGAGCTC 700
340 euGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLysGly 356
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
701 TGGAGTGGGTATTCCTGAAATGGCAAGGAGGCATGCCCTTGACAAAGGT 750
357 GluAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleVa 373
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
751 GAGCGAGTTAAATTTTGAAGGTGCAGTTGTGACAGCAGATCGAATTGT 800
373 lThrValSerGlnGlyTyrrSerTrpGluValThrThrAlaGluGlyGly 390
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
801 GACTGTCAAGTAAAGGTATTTCATGGAGGTCCAACTGCTGAAGGTGGAC 850
390 lnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIle 406
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
851 AGGGTCTCAATGAGCTCTTAAGCTCCCGAAAGAGTGTATTAAACGGAATT 900
407 ValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLe 423
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
901 GTAATGGAATGACATTAATGATTGGAACCTTCGACGGACAAATGTAT 950
423 uProHisHisTyrrSerValAspAspLeuSerGlyLysAlaLysCysLysA 440
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
951 CCCTTGTGCTATTCTGTTGATGACCTCTCGAAAGGCCAAATGTAAANA 1000
440 laGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeu 456
::: ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1001 GTGCATTGCAGAGGAGCTGGGTTTACCTATAAGGCCTGAAGTTCTCTCTG 1050
```

```
457 ILeGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleY 473
|||||
1051 ATTGGTTTATTGAGAGATTGGATTATCAGAAAGGCAATGATCTCATCA 1100
|||||
473 sMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuG 490
|||||
1101 ACTTATCATACCATCTCTATCGGAGCAGCGTTCAATTTGTCTGCTTG 1150
|||||
490 LySerGlyAspProIlePheGluGlyTyrMetArgSerThrGluSerSer 506
|||||
1151 GATCTGCTGACCCAGAGCTCGAAGACTGGATGAGATCTACAGATCGGAC 1200
|||||
507 TyrLysAspLysPheArgGlyTyrValGlyPheSerValProValSerHi 523
|||||
1201 TTCAGGATAAATTTCTGGATGGGTTGGATTTAGTTTCCAGTTTCCCA 1250
|||||
523 sArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluP 540
|||||
1251 CCGAATACTCGCGCTGGCATATATTGTAATGCCATCCAGATTCCGAAC 1300
|||||
540 roCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVal 556
|||||
1301 CTTGTGCTCAATCAGCTATATGCTATGTCAGTATGGCACAGTTCTCT 1350
|||||
557 ValHisGlyThrGlyLeuArgAspThrValGluThrPheAsnProPh 573
|||||
1351 GTCCATGCAACTGGGGGCTTAGAGATACTGTGGAGAACTTCAACCCTTT 1400
|||||
573 eGlyAlaLysGlyGluGlyThrGlyTyrAlaPheSerProLeuThrV 590
|||||
1401 CGGTGAGATGAGAACAGGTACAGGTGGGATTCGACCCCTTAACCA 1450
|||||
590 aLaspLysMet..LeuTrpAlaLeuArgThrAlaMetSerThrPheArgG 606
|||||
1451 CGGAACAACATGTTGTGGACATTTGGAACACTGCAATTTCCACATACA 1500
|||||
606 uHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAsp 623
|||||
1501 GCACAAATCTCTCTGGGAGGCTCATGAAGGGGCGCATGTCAAAAGACT 1550
|||||
623 IsThrTrpAspHisAlaPro..SerSerThrSerArgSerSerGlyPr 639
|||||
1551 TCACGTGGGACCATGCCCTGAACAATACGAACAATCTTCCAGTGGGCC 1600
|||||
639 oSerTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLys 656
|||||
1601 TTCATCGATCGACCCCTATGTCATGTAAGGGGAC..... 1634
|||||
656 ysGlu..SerProSerAlaLeuLysThr...SerSerSerPheArgG 671
|||||
1635 ..CACAGTGTGTCTCTTAAAGATCTTCAGTTCTTCATCCTACAGTGG 1682
|||||
671 yProGlu..GlyTyrProCysThrLeuArgCysProAlaThrValGluSer 687
|||||
1683 AAGGATGAGAAACCCCTGTACATTACAT..... 1711
|||||
688 GlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAspGlyCy 704
|||||
1712 .....GGAAGGCGGACC.....GGCTG 1728
|||||
704 sAlaAlaAlaAlaValThrAlaSerGlyArgGlnLeuGlnPheTrpG 721
|||||
1729 TTGGGTGGCTCCAATGTCCAGGTAGCCTTACGGAGCCTCTGCTTTGGC 1778
|||||
721 lyIleArgLysGlyCysAlaAlaGly 729
|||||
1779 TGGTT.....TGTCCTCAAGGA 1795
|||||
```

seq_name: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:US-09-625-406-20

seq_documentation_block:

; Sequence 20, Application US/09625406

; GENERAL INFORMATION:

```
APPLICANT: Keeling, Peter
APPLICANT: Guan, Hanping
TITLE OF INVENTION: Starch Encapsulation
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/625,406
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/941,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-09-625-406-20
```

```
alignment_scores:
Quality: 2546.50      Length: 538
Ratio: 4.935          Gaps: 3
Percent Similarity: 95.911 Percent Identity: 87.918

alignment_block:
US-09-674-824-2 x US-09-625-406-20 ..

Align seg 1/1 to: US-09-625-406-20 from: 1 to: 1620

95 ValGlyGluLeuAla...ProAspLeuLeuLeu.....GluGlyIleAl 108
|||||
4 GTCCGGAGCTGAGCAGGAGGAGCCCTCGGTCTCGAACCTGAAGGGATTGC 53
|||||

108 aGluAspSerIleAspSerIleValAlaAlaSerGluGlnAspSerG 125
|||||
54 TGAAGTTTCCATCGATAACACACAGTAGTTGTGGCAAGTGAGCAAGATTCTG 103
|||||

125 luIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSerIle 141
|||||
104 AGATTCTGTTGGAAAGGAGCAGCTCGAGCTAAAGTAACAAACCAATT 153
|||||

142 ValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuG 158
|||||
154 GTCTTTGTAAACCGCGAAGCTTCTCTTATGCAAGCTTGGGGGTCTAGG 203
|||||

158 yAspValCysGlySerLeuProIleAlaAlaAlaArgGlyHisArgV 175
|||||
204 AGATGTTTGTGGTTCAATGCCAGTTGCTTCTGCTGCTGTCACCGTGS 253
|||||
```

```
175 alMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyr 191
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 TGATGGTGTAAATGCCAGATATTTAAATGGTACCTCCGATAAAGATAT 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
192 AlAlysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGI 208
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 GCAATGGATTTTACACAGAAACACATTCGGATTCCATGCTTTGGCGG 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
208 ySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrpv 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 TGAACATCAAGTACCTTCTCCATGAGTATAGAGATTCACTTGACTGG 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
225 alPheValAspHisProSerTyrHisArgProGlySerLeuTyrGlyVal 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 TCTTTGTTGATCATCCCTCATATCACAGACCTGGAAATTTATATGGAG 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
242 AsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTy 258
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
454 RAGTTTGGTCTTTGGTGATTAATCAGTTTCAGATACACACTCCTTGTG 503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
258 rAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrG 275
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
504 TGCTGCATGTGAGGCTCCTTTGATCCTTGAATTTGGAGGATATATTTAT 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
275 LyGlnAsnCysMetPheValValAsnAspTyrPheHisAlaSerLeuValPro 291
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
554 GACAGAAATTCATGTTGTTGTCATGATTTGGCATGCCAGTCTAGTGCCA 603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
292 ValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerAr 308
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
604 GTCTTCTTGTGCGAAATATAGACCATATGTTGTTTATTAAGACTCCCG 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
308 gSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlaAs 325
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
654 CAGCATTCCTGTAATACATAATTTAGCACATCAGGGGTGAGAGCCTGCA 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
325 exThrTyrProAspLeuGlyLeuProGluTyrTrpTyrGlyAlaLeuGlu 341
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
704 GCACATATCTCACCTTGGTTGCCACCTGTAATGTTATGGAGCTCTGGAG 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
342 TrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyGluAl 358
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
754 TGGGTATCCCTGAAATGGCGAGGAGCATGCCCTTGACAAAGGTGAGGC 803
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
358 aValAsnPheLeuLysGlyAlaValThrAlaAspArgIleValThrV 375
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
804 AGTTAATTTTGAAGGTGCAGTTGTGACAGCAGATCGAATCGTGACTG 853
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
375 alserGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGlnGly 391
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
854 TCAGTAAGGGTTATTCGTGGAGGTCCAACTGCTGAAGGTGGACAGGC 903
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
392 LeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAs 408
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904 CTCATGAGCTCTTACGTCCAGAAAGATGTATTAAACGGAATGTATAA 953
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
408 nGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuProH 425
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
954 TGAATTTGACATTAAATGATTGGAAACCTGCCACAGACAAATGTATCCCT 1003
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
425 lsHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGlu 441
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1004 GTCATTTATCTGTGTGATGACCTCTCGAAAGGCCAAATGTAAGGTGCA 1053
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleGI 458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1054 TTCAGAGAGAGCTGGGTTTACCTATTAAGGCTGATGTTCTCTGATTGG 1103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetA 475
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1104 CTTTATTGGAAGGTTGGATTATCAGAAAGGCATTGATCTCATTCAACTTA 1153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
475 laIleProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1154 TCATACCAGATCTCATCGGGAAGATGTTCAATTTGTCATGCTTGGATCT 1203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
492 GlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLy 508
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1204 GGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAA 1253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
508 sAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArgI 525
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1254 GGATAAATTTCTGGTGGTGGTGGATTAGTTTCCAGTTTCCACCCGAA 1303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
525 leThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCys 541
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1304 TAACTGCCGCTCGATATATTTGTTAATGCCATCCAGATTTCGAACCTTGT 1353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValH 558
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1354 GGTCTCAATCAGCTATATGCTATGCAGTATGGCACAGTTCCTGTTGCCA 1403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1404 TGCAACTGGGGCCTTAGAGATACCGTGAGAACTTCAACCTTTCGGTG 1453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
575 laLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValAsp 591
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1454 AGAATGGAGAGCAGGGTACAGGGTGGGCATTCGCCACCCCTAACCCAGAA 1503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
592 LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisL 608
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1504 AACATGTTTGGGACATTCGCACTGCAATATCTACATACAGGGAACACA 1553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
608 ysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 624
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1554 AGTCTCTCTGGGAAGGCTAATGAAGCGAGGCATGTCAAAAGACTTCACG 1603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
625 TrpAspHisAla 628
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1604 TGGGACCATGCC 1615
```

seq_name: /cgn2_6/ptodata/1/pna/us093_COMB.seq.us-09-345-214-6

seq_documentation_block:

; Sequence 6. Application US/09345214
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: Zea mays
US-09-345-214-6

alignment_scores:
Quality: 2495.00 Length: 508
Ratio: 5.051 Gaps: 1
Percent Similarity: 97.244 Percent Identity: 89.567

alignment_block:

US-09-674-824-2 x US-09-345-214-6/rev ..
Align seg 1/1 to reverse of: US-09-345-214-6 from: 1 to: 1528

134 GlnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPr 150
:::CGAGCTAAAGTAACACAAGCAATTCCTTTGTGAACGGCGAAGCTTCCC

1527 oTyRAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleA 167
1477 TTATGCAAAGCTCGGGGCTTAGGAGATGTTTGTTGGTTCATTGCCAGTGTG 1428

167 IalEuAlaAlaArgGlyHisArgValMetValMetProArgTyrLeu 183
1427 CPTCTGCTGCTCGTGCCTCACCGTGTGATGGTGTGAATGCCAGATAATTA 1378

184 AsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHi 200
1377 AATGGTACCCTCCGATAGAATATATGCAAAATGCATTTTTTACACAGAAAACA 1328

200 sIleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
1327 CATTCGGATTCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATG 1278

217 LuTyrArGAspAsnValAspTrpValPheValAspHisProSerTyrHis 233
1277 AGTATAGAGATTACGTTGCACGGGTGTTGTGTGATCATCCTCATATCAC 1228

234 ArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGl 250
1227 ACACCTCGAAATTTATATGGAGATAAGTTGGTGCTTTTGGTGAATACA 1178

250 nPheArgTyrThrLeuLeuCystYrAlaAlaCysGluAlaProLeuIleL 267
1177 GTTCAGATTACACTCCTTCTGATGTGATGTGAGGCTCTTTGTATCC 1128

267 euGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsn 283
1127 TTGNAATGGGAGGATATATTTATGCACAGAATTGCATGTTGTGTGCAAT 1078

284 AspTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPr 300
1077 GATTGGCATGCCAGCTAGTGCCAGTCCTTCTTGTCGCAAAATATAGACC 1028

300 oTyRGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeua 317
1027 ATATGGTGTTTATAAGACTCCCGACGATCTTGTAATAACATAAATTTAG 978

317 IaHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 333
977 CACATCAGGGTGTAGAGCTGCAAGCACATATCCTGCACCTGGGTTGCCA 928

334 ProGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgAr 350
927 CCTGAATGGTATGAGCTCTGGAGTGGGATATCCCTCAATGGCGAGGAG 878

350 gHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValV 367
877 GCATGCCCTTGACAAGGCTGAGGCAGTTAATTTTTTGAAGGTGCAGTTG 828

367 alThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluVal 383
827 TGACAGCAGATCGAATCGTCACTGTCAAGTGGGTTATTCGTGGGAGGTC 778

384 ThrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLy 400
777 ACAACTGCTGAAGGTGCACAGGCCCTCAATGAGCTCTTAAGCTCCAGAAA 728

400 sSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
727 GAGTGTATTAAACGGAAATTTGAATGGAAATTCACATTAATGATGTGAACC 678

417 roThrThrAspLysCysLeuProHisHistySerValAspAspLeuSer 433
677 CTGCCACAGACAAATGTATCCCCCTGTCTATTATTCCTGTGATGACCTCTCT 628

434 GlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVa 450

[illegible]

630 GGCTCAAGTCAGCAGGATCTTGAGATCATGGATGCCAATGAGCAACCTC 679
134 lnaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
680 AAGCTAAAGTTACACGTAGCATCGTTGTGTGACTGGTGAAGCTGCCTC 729
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
730 TATGCAAAAGTCAGGGGGTGGAGATGTTGTGCTGCTTACCAATGTC 779
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
780 TCTTGCTCTCGTGCACCCAGTCAGTGGTGTAAATGCCAAGATACTTAA 829
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
830 ATGGGTCTCTCATAAAAACATATGCCAAGGCATATATACATGCCAAGC 879
201 ileLysIleProCysPheGlyCysSerHisGluValThrPhePheHisG1 217
880 APTAAGATTCATGCTTTGGGGATCATGAAGTGACCTTTTTCATGA 929
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
930 GTATAGAGACAACGTCGATTTGGGTGTTTGTGATCATCCGTATATACA 979
234 rProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
980 GACCAGGAAGTTATATGAGATAATTTGGTGCTTTTGGTGATTAATCAG 1029
251 PheArgTyrThrLeuLeuCystyrAlaAlaLacysGluAlaProLeuIleLe 267
1030 TTCAGATACACACTCTTTGCTATGCTGCATCGAGGCCCTTAATCCT 1079
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
1080 TGAATTTGGGAGGATATATTTATGGACAGAAATTCATGTTGTGTGAAGC 1129
284 sPTPrHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
1130 ATTGGCATGCCAGCTTGTGCCAGTCTCTTGTGTCGCAAAATATAGACCA 1179
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
1180 TACGGTGTTTACAGAGATCCCGCAGCACCTTGTATATATAAATTTAGC 1229
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
1230 ACATCAGGGTGTGGAGCCTGCAAGTACATATCTGATCTGGGATGCCCTC 1279
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
1280 CTGAATGGTATGGAGCTTTAGATGGTATTTCCAGAATGGGCAAGGAGG 1329
351 HisAlaLeuAspLysGlyGluAlaValAlaAsnPheLeuLysGlyAlaValVa 367
1330 CATGCCCTTGACAAGGCTGAGGCAGTTAACTTTTGAAGGAGCAGTGT 1379
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
1380 GACAGCAGATCGGATGTGACCTCAGTCAGGGTTATTCATGGAGGTCA 1429
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1430 CAACTGTGTAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAAAA 1479
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
1480 AGTGATGTGAATGGAATGTAAATGGAAATTTGACATTAATGATTGGAAACC 1529
417 oThrThrAspLysCysLeuProHisIstYrSerValAspAspLeuSerg 434
1530 CACCACAGACAAGTGTCTCCCTCATCATATTTCTGTCGATGACCTCTCTG 1579

434 lYlysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1580 GAAAGCCAAATGTAAGCTGAATTCAGAAGGAGTTGGGTTTACCTGTA 1629
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1630 AGGGAGGATGTTCTCTGTGATTTGGTATTTATTTGGAAAGACTGGATTACCAGAA 1679
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
1680 AGGCATGTGATCTCATTTAAATGCCATTCAGAGCTCATGAGGAGGACG 1729
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
1730 TGCAATTTGTTCATGCTTGGATCTGGGATCCCAATTTTGAAGGCTGGATG 1779
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
1780 AGATCTACCGAGTCGAGTTACAAGGATAAATTCGTGGATGGGTGGATT 1829
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1830 TAGTGTCCAGTTTCCACAGATAAATTCGAGGTTGCGATATATTTGTAA 1879
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1880 TGCCATCGAGATTTGAACCTTGGGCTCTTAATCAGCTATATGCTATATGCAA 1929
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
1930 TATGTCAGATTCCTGTAGTTCATGGAACTGGGGCCCTCCGAGACACAGT 1979
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
1980 CGAGACCTTCAACCTTTTGGTGCAAAAGGAGAGAGGGGTACAGGGTGG 2029
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
2030 CGTTTCAACCTTCAACCTTGGTGCAAAAGGAGAGGGGTACAGGGTGG 2079
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
2080 ATGTCGACATTCAGGAGCACAAAGCCGCTCTGGAGGGGCTCATGAAGCG 2129
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSerA 634
2130 AGGCATGACGAAAGACCATAGCTGGGACCCTATGCCCCGAGCAGTCCCTTC 2179
634 rGSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650
2180 GATCTTCAGTGGGCTTCTGGTGGACCAACCTACGTCATGTAGAGGGGA 2229
651 LeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSe 667
2230 CTGGGAGGTCCCAAGTCGAGTCTCTTTCAGCTCTGAAGACATCCTCTTC 2279
667 rSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAlat 684
2280 ATCTCTCCGGCCCGGAAAGGATACCCCTGTACATTGGCTGTCTCTGCTA 2329
684 hrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThr 700
2330 CAGTAGAGTCGCAATCGCCTCTTGTGCTTGTGCTTGTGCTGAGAGAA 2379
701 TyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnle 717
2380 TATGAGGCTGTGCTGCTCGCGGTGACAGCTTCGGGTGGACGACAGTT 2429
717 uGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaL 734
2430 ACAGTTTTGGGAAATAAGAAAGGATGTCTGCTCAGGATGGTTAACAGCAA 2479

734 yshishisSerAspGlySerLeuSerValArgValThrAlaGluIleArg 750
 2480 AGCACCACTCAGATGCGACCTCTCTGTCGGTGTACAGCTGAATCAGA 2529
 751 AsnGlnLeuValThrLeu 756
 2530 AACCAACTGGTGACTCTT 2547

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:US-09-508-377-11

seq_documentation_block:
 ; Sequence 11, Application US/09508377
 ; GENERAL INFORMATION:
 ; APPLICANT: KALEEN, ZHONGYILI
 ; APPLICANT: MORELL, MATTHEW
 ; APPLICANT: RAHMAN, SADJOUR
 ; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS
 ; FILE REFERENCE: 054270/0126
 ; CURRENT APPLICATION NUMBER: US/09/508, 377
 ; PRIOR FILING DATE: 2000-06-09
 ; PRIOR APPLICATION NUMBER: AU PP 2509
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: PCT/AU98/00743
 ; PRIOR FILING DATE: 1998-09-11
 ; PRIOR APPLICATION NUMBER: AU PP 9108
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 2662
 ; TYPE: DNA
 ; ORGANISM: Triticum tauschii
 US-09-508-377-11

alignment_scores:
 Quality: 3930.00 Length: 757
 Ratio: 5.247 Gaps: 1
 Percent Similarity: 98.943 Percent Identity: 98.415

alignment_block:

US-09-674-824-2 x US-09-508-377-11 ..

Align seg 1/1 to: US-09-508-377-11 from: 1 to: 2662

1 MetaAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
 257 ATGCGCGGACGGCGTGGCGCGGGTGCCTCGCCCGACGCTCGCCCT 306
 17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
 307 CGCGCGCGATCCGGCGACGGCGCGCGGGGTCCGCTGGCTGCTCGCG 356
 34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
 357 CGCGGCTCCGGCGCTTGGCGGGGGCGGTACCTTCCGAGCTCAGCAGG 406
 51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67
 407 GAGGCT 456
 67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
 457 CGTGGCAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506
 84 laProThrGlnProProProProAlaGlyValGlyGluLeuAlaPro 100
 507 CCGCGACG 556
 101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVal 117
 557 GACCTCTCTCGAAGGATTCCTGAGGATTCATCGACGACATTAATGT 606

117 lalaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
 607 GGCTGCAAGTCAGCAGGATTCAGATCATGGATCGAATGCAACCTC 656
 134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
 657 AAGCTAAAGTTACAGTACGATCGTGTGTGACTGGTGAAGCTGCTCCT 706
 151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
 707 TATGCAAGTCAGGCGGCTGGGAGATGTTGTGGTTCGTTACCAATGC 756
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
 757 TCTTGTCTGCTCGTGTCCACCGCTGTGATGGTTGTAATGCCAAGATCTGA 806
 184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
 807 ATGGTCTCTGATAAAACTATGCAAAAGCATATACACTGGGAAGCAC 856
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217
 857 ATTAAGATTCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGA 906
 217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
 907 GTATAGAGACACGTCGATTTGGGTGTTGTGCGATCATCCGTCATATCA 956
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
 957 GACCAGGAATTTATATGAGATAAATTTGGTCTTTTGGTGATAATCAG 1006
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
 1007 TTCAGATACACACTCTCTTGTATGCTGATGCGAGGCGCCCACTAATCCT 1056
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
 1057 TGAATTTGGGAGGATATTTATGGACAGAAATTCATGTTGTGTGTAAGC 1106
 284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
 1107 ATTGGCATCCAGCCTTGTGCCAGTCTCTTCTGCTGCAAAATATAGACCA 1156
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
 1157 TACGGTGTTCAGAGATTCGCCGACGACCTTGTATATACATAAATTTAGC 1206
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
 1207 ACATCAGGGTCTGGAGCCTTGCAAGTACATATCTCATCTGGGATTCACAC 1256
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
 1257 CTGAATGGTATGGAGCTTTAGAATGGGTATTTCCAGAAATGGGCAAGGAGG 1306
 351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
 1307 CATGCCCTTGACAAGGGTGAGCAGCTTAACCTTTTGAAGGAGGAGCTCGT 1356
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
 1357 GACAGCAGATCGAATTTGACCGTCAGTCAGGTTATTTCATGGGAGGTCA 1406
 384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
 1407 CACTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAA 1456
 401 SerValLeuAsnGlyIleValIasnGlyIleAspIleAsnAspTrpAsnPr 417
 1457 AGTGATTGAATGGAATTTGTAATGGAATTCACATTAATGATTCGAACCC 1506
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434

|||||
1507 CACCACAGCAAGTGTCCCTCATATTCTGTCGATGACCTCTCTG 1556
434 lYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1557 GAAAGGCCAANTGAAGCTGANTGCAGAGGAGCTGGGTTTACCTGTA 1606
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1607 AGGAGGATCTCTCTCGATTGGCTTTATTGGAAGACTGGATTACAGAA 1656
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
1657 AGGATTGATCTCATTAATAATGCCATCTCCAGAGCTCATGAGGAGGACG 1706
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
1707 TGCAGTTTGTTCATGCTTGGATCTGGGATCCAAATTTTGAAGCTGGATG 1756
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
1757 AGATCTACCGAGTCGATTAACAAGGATAAATCCGTGGATGGGTGGATT 1806
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1807 TAGTGTCCAGTTTCCACAGAAATACTGCAGGTTCGATATATTGTTAA 1856
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1857 TGCCATCCAGCTTTGAACCTTGTGCTTAATATCATCATATATGCTATGCAA 1906
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
1907 TATGTACAGTTCTGTAGTTTCATGGAACTGGGGCTCCGAGACACAGT 1956
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584
1957 CGAGACCTTCAACCTTTTGTGTGCAAAAGGAGAGGAGGTACAGGGTGGG 2006
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
2007 CGTCTCACCGCTAACCGTGCAAGATGTTGTGGGCATTCGGAACCCG 2056
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
2057 ATGTCGACATTCAGGAGCACAAAGCGCTCTGGAGGGCTCATGAAGCG 2106
617 gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer 633
2107 AGGCATGACGAAAGACCATACGTGGGACCATGCCCGCGAGCAGTACGAGC 2156
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
2157 AGATCTTCGAATGGCGCTTCGTGGACCAACCTACGTCATGTAGACGGG 2206
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
2207 ACTGGGAGGTGCAAGCGCGGCTCTCTGAGCTCTGAAGACATGTTCT 2256
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683
2257 CATCTCTCCGCGCGCGGAAGGATACCCCTGTACATTGGCTTCTCTGCT 2306
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700
2307 ACAGTAGATCGCAATGCGCTCTGCT.TGGTCCGCGGTTCCGAGAT 2355
700 rTyrAspGlyCysAlaAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717
2356 AGATGACCGCTGCTGCTGCGCGGCTGACAGCTTCGGGTGATGACAT 2405
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733
|||||

2406 TACAGTTTTGGGAATAAGGAAGGATGTGTCGAGGATGGTTAACAGCA 2455
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750
2456 AAGCACCACTCAGATGGCAGCCTCTCTCTCCGTGTACAGCTGAATCAG 2505
750 gAsnGlnLeuValThrLeu 756
2506 AAACCAACTGGTGACTCTT 2524
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:US-09-952-677-1
seq_documentation_block:
; Sequence 1, Application US/09952677
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Lorz, Horst
; Lutticke, Stephanie
; Walter, Lennart
; Froberg, Claus
; Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; FROM WHEAT WHICH ARE INVOLVED IN STARCH
; SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/09/952,677
; FILING DATE: 14-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,390
; FILING DATE: 19-Nov-1998
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum L.
; STRAIN: cv. Florida
; HAPLOTYPE: ca. 21 d Caryopses
; IMMEDIATE SOURCE:
; LIBRARY: cDNA library in pBluescript sk (-)
; CLONE: TaSSS
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: 3..2017
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-952-677-1

alignment_scores:
  Quality: 3612.00      Length: 671
  Ratio: 5.383          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-674-824-2 x US-09-952-677-1
..
Align seg 1/1 to: US-09-952-677-1 from: 1 to: 2239

      86 ThrGlnProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102
      3  ACGCAGCGCCCTTGGCGGACCGCGGTGGGGAACTCGCGCCGACCT 52
102  uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaA 119
      53  CCTGCTCGAAGGGATTGCTGAGGATTCATCGACAGCATATTGTGCTG 102
119  laSerGluGlnAspSerGluIleMetAspAlaAsnGluInProGlnAla 135
      103  CAAGTGACGAGGATTCAGATCATGATGCGAATGAGCAACCTCAAGCT 152
136  LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152
      153  AAAGTTACAGCTAGCATCGTGTGTGCTGCTGGTGAAGCTGCTCTTAATG 202
152  aLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuA 169
      203  AAAGTCAGGGGGGTGGGAGATGTTTGTGCTTGTACCAATTGCTCTTG 252
169  laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185
      253  CTGCTCGTGTCCAGGAGTGATGTTGTAATGCCAAGATACCTTAAATGG 302
186  SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleTy 202
      303  TCCTCTGATAAAACATATGCAAGGCATTATACACTCGGAAGCACATTAA 352
202  sIleProCysPheGlyCysHisGluValThrPhePheHisGluTyrA 219
      353  GATTTCATGCTTTGGGGGATACATGAAGTGACCTTTTTCATGAGTATA 402
219  rGAspAsnValAspTyrValPheValAspHisProSerTyrHisArgPro 235
      403  GAGACAACGTCGATTGGGTGTTTGTGATCATCCGTCATATACAGACCA 452
236  GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252
      453  GGAAGTTTATATGGAGATAATTTTGTGCTTTTGGTGATAATCAGTTTCA 502
252  gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269
      503  ATACACTCCCTTGTATGTCATGCGAGGCGCCACCTAAATCCTTTGAAT 552
269  euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTyr 285
      553  TGGGAGGATATATTTATGACAGAAATTCGATGTTGTGTGAACGATTGG 602
286  HisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGl 302
      603  CATGGCAGCCTTGTGCGAGCTCTCTTGTGCAAAATATAGACCATACGG 652
302  yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
      653  TGTTTACAGAGATTCCCGCAGCACCTTGTATACATAATTTAGCACATC 702
319  InGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProProGlu 335
```

```
|||||
703  AGGCTGTGGAGCCTGCAAGTACATATCTGATCTGGGATTGCCCTCTGAA 752
336  TrpTyrGlyAlaLeuGluTyrValPheProGluTyrPalaArgHisAl 352
|||||
753  TGGTATGGAGCTTTAGAAATGGGTATTTCAGAAATGGCAAGGAGGATGC 802
352  aLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrA 369
803  CCTTGACAAGGGTGAGGACGTTAACTTTTGAAGGAGGAGCTGTGACAG 852
369  laAspArgIleValThrValSerGlnGlyTyrSerTyrGluValThrThr 385
853  CAGATCGGATTGTGACCGTCAGTCAGGGTTATTTCATGGAGGTCACAACT 902
386  AlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402
903  GCTGAAGGTGACAGCGCCTCAATGAGCTCTTAAGCTCCCGCAAAAGTGT 952
402  LeuAsnGlyIleValAsnGlyIleAspIleAsnAspTyrAsnProThrT 419
953  ATTCAATGGAAATTGTAATGGAATTGACATTAATGATTGGAACCCACCA 1002
419  hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435
1003  CACACAAGTCTCCCTCATCATATTCTGTCGATGACCTCTCTGGAAG 1052
436  AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgL 452
1053  GCCAAATGTAAGCTGAATTCAGAAAGGAGTTCGGTTTACCTGTGAAGGA 1102
452  uAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469
1103  GGATGTTCTCTGATTGGGCTTTATTGGAAGACTGGATTACCGAGAAAGCA 1152
469  leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485
1153  TTGATCTCATTAATAATGGCCATTCAGAGCTCATGAGGGAGGACGTGCAA 1202
486  PheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMetArgSe 502
1203  TTTGTCATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATGAGATC 1252
502  rThrGluSerSerTyrLysAspLysPheArgGlyTyrValGlyPheSerV 519
1253  TACCGAGTCGAGTTACAAGATAAATTCCTGGATGGGTGGATTAGTGTG 1302
519  alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535
1303  TTCCAGTTTCCACAGATAAATTCAGAGTTTCGATATATTGTTAATGCCA 1352
536  SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGl 552
1353  TCGAGATTTGAACCTTGGCGTCTTAATCAGCTTATATGCTATGCAATAGG 1402
552  yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569
1403  TACAGTTCTCTGTAGTTTCATGGAACCTGGGGCTCCGACACACAGTCGAGA 1452
569  hrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTyrPalaPhe 585
1453  CTTTCAACCCCTTTTGGTGCAAAGGAGAGGAGGTACAGGGTGGCGGCTTC 1502
586  SerProLeuThrValAspLysMetLeuTyrPalaLeuArgThrAlaMetSe 602
1503  TCACCGCTTAACCGTGGCAAGATGTTGTGGGCATTTGGCAACCGCATGTC 1552
602  rThrPheArgGluHisLysProSerTyrPalaGlyLeuMetLysArgGlyM 619
1553  GACATTCAGGAGGACACAAGCCGCTCTGGAGGGGCTCATGAAGCGAGGCA 1602
619  etThrLysAspHisThrTyrPalaPheHisAlaProSerSerThrSerArgSer 635
|||||
```

```
1603 TGACGAAAGACCATACGTGGACCATGCCCGAGCAGTAGTACGACGATCT 1652
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuGl 552
1653 TCGAGTGGGCGCTTCGTGGACCAACCCATAGTCATGTAGCGGAGCTGGG 1702
652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerP 669
1703 GAGGTCCAAGTGGGAGTCTCCTTCAGCTCTGAAGACATCTCTTCATCT 1752
669 heArgGlyProGluGlyTyProCysThrLeuArgCysProAlaThrVal 685
1753 TCCGCGCGCGGAGGATACCCCTGTACATTGCTGTCTGTCTACAGTA 1802
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrYrAs 702
1803 GAGTCCGAATGCGCCTGCTTGTCTTGTGTCGCGGTTTCGAGAATATGA 1852
702 pGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnP 719
1853 CGGCTGTGCTGTCGCGCGGTGACAGCTTCGGGTGGACGACATTACAGT 1902
719 heTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
1903 TTTGGGCAATAAGGAAGGATGCTGCTCAGGATGGTTAACAGCAAGCAC 1952
736 HisSerAspGlySerLeuSerValatqValThrAlaGluIleArgAsnGl 752
1953 CACTCAGATGCGACGCTCTGTCGGTGTACAGCTGAAATCAGAAACCA 2002
752 nLeuValThrLeu 756
2003 ACTGGTGACTCTT 2015

seq_name: /cn2_5/ptodata/2/pna/us09_NEW_COMB.seq2:us-09-931-297-1

seq_documentation_block:
; Sequence 1, Application US/09931297
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; starch synchases from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/931,297
; FILING DATE: 16-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/192,909
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 1:
```

```
SEQUENCE CHARACTERISTICS:
LENGTH: 2383 base pairs
TYPE: nucleotide
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Zea mays
TISSUE TYPE: endosperm
FEATURE:
NAME/KEY: CDS
LOCATION: 2..1950
OTHER INFORMATION: /function= "starch synthesis"
/product= "soluble starch synthase"
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-931-297-1

alignment_scores:
Quality: 2809.50 Length: 749
Ratio: 4.411 Gaps: 10
Percent Similarity: 85.047 Percent Identity: 73.565

alignment_block:
US-09-674-824-2 x US-09-931-297-1 ..
Align seg 1/1 to: US-09-931-297-1 from: 1 to: 2383
4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuAlaAs 20
41 TCGCGCGTGGCGCGCGTCCCTCCTC..... 67
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
68 CTCGCGCGCGCGCGCTGGCGCGCGCGCTCGCGCACCGCGCGCGCGC 116
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyValAlaGluLeuSerArg 50
117 GCGCGGTCCACGCGCTGCTCGCGCGCGCGCTCGCGGAGCTGACGAGG 166
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
167 GAGGCGCGCGCGCGCGCGCGATGCCACCGCGCTGCTGCGCGCGCGCT 216
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
217 CGTGCGCGGCTTCCTCGCG...CGCGCGCGCGCGCGCGGTGAGCGCG 263
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
264 CATTGACGCGCGCGCGCGCGCGCGCGCGCGCTGGGGGCTCCTCGGTC 313
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
314 GAA.....CCTGAAGGATTGCTGAAGTTCCATCATCAACACAGTAGT 357
117 lAlaAlaSerGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
358 TGTGGCAAGTAGCAAGATTCTGAGATTGTGTTGGAAGAGGAGCAAGCTC 407
134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
408 GAGCTAAGTAACACAAACATTGCTTTGTAAGTGGGAAGCTTCCTCT 457
151 TyrAlaLysSerGlyLeuGlyAspValCysGlySerLeuProIleAl 167
458 TATGCAAGTCTGGGGGTCTAGAGATGTTTGTGTTTCATTGCGCAGTTGC 507
167 aleuAlaAlaArgGlyHisArgValMetValValMetProArgTyLeuA 184
508 TCTTGCTGCTGCTGCTGCTACCGCTGATGTTGTTAATGCCAGATATTAA 557
```

184 snGlySerSerAspLysAsnThrAlaLysAlaLeuThrAlaLysHis 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
558 ATGGTACCTCCGATAGAAATATGCAATGCAATTTACACAGAAAACAC 607
201 IleLysIleProCysPheGlySerHisGluValThrPhePheHisG1 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
608 ATTCCGATTCCCATGCTTTGGCGGTGAACATGAAGTTACCTCTTCCATGA 657
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerThrHisA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
658 GTATAGAGATTCCAGTTGACTGGGTGTTTGTGATCATCCCTCATATCACA 707
234 rgProGlySerLeuThrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
708 GACCTGGAAATTTATATGAGATAAGTTTGGTGCTTTTGGTGATNAATCAG 757
251 PheArgThrThrLeuLeuCystyrAlaAlaCysGluAlaProLeuIleLe 267
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
758 TTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGGTCT 807
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
808 TGAATTGGGAGATATATTTATGGACAGAATTGCATGTTTGTGCAATG 857
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
858 ATTGGCATGCCAGTCTAGTCCAGTCTTCTTCTGCGCAAAATATAGACCA 907
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
908 TATGTTGTTTAAAGACTCCCGCAGCATCTTGTGTAATACATAATTAGC 957
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
958 ACATCAGGGTGTAGAGCTCGCAACACATATCTCGACCTTGGGTGGCCAC 1007
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1008 CTGAATGGTATGGAGCTCGAGTGGGTATTCCTGTAATGGCGAGGAGG 1057
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1058 CATCCCTCTGCAGAGGTGAGCGAGTAAATTTTGAAGGTGCAGTTGT 1107
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1108 GACAGCAGATCGAATCTGTACTGTCACTAAGGGTTATTCTATGGAGGTCA 1157
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1158 CAACCTGCTGAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1207
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1208 AGTGATTAAACGAATGTGAATGGAATGACATTAATGATTGGAACCC 1257
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspLeuSerG 434
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1258 TGGCAGACAGAAATGTATCCCTGTCTATTATTCTGTGATGACCTCTCTG 1307
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1308 GAAAGGCCAANTGAAGGTGCATTCGAAAGGAGCTGGGTATTACCTATA 1357
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLys 467
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1358 AGGCCTGATGTTCTCTGATGGCTTTATTGGAAGATTGGATTATCAGAA 1407
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1408 AGGCATTGATCTCAATCACTTATATCACCAGATCTCATCGCGGAAGATG 1457

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1458 TTCAATTTGTCTATGCTTGGATCTGGTGAGCCAGAGCTTTGAAGATTGGATG 1507
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1508 AGATCTACAGAGTCTATCTTCAAGGATAAATTTCTGGATGGTGGATT 1557
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1558 TAGTGTTCACAGTTTCCACCACTGCTGCTCAATCAGCTATATGCTATGCAG 1607
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1608 TGCCATCCAGATTCCAGCTTGTGCTCAATCAGCTATATGCTATGTTAA 1607
551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1658 TATGGCAGACTTCTGTGTCATGCAACTGGGGGCTTAGAGATACCGT 1707
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1708 GGAGAACTTCAACCTTTTCGGTGAGAAATGGAGACGAGGTACAGGTGGG 1757
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1758 CATTCGACCCCTTAACACAGAAACATGTTGTGGACATTGCGAACTGCA 1807
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1808 ATATCTACATACAGGGAACACAGCTCTCTCTGGGAGGGCTAATGAAGCG 1857
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1858 AGGCATGTCAAAAGACTTTCAGTGGGACCATGCCGTGAAACAATACGAAC 1907
634 ArgSerSerGlyProSerTrpTrpAsnProThrSerCysArgArgG1 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1908 AAATCTTCCAGTGGGCTTCTATCGATCGACCCCTATGTCATGTAAAAAGG 1957
650 yLeuGlyArgSerLysCys.GluSerProSerAlaLeuLysThrSerSer 666
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1958 A.....CCAAAGTGGTGGTTCCTTGAAGATCATCAGTTCATCATCC 1998
667 SerSerPheArgGlyProGluGlyTyrPro.....CysThrLeuArgCy 681
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1999 TATAGTAAGCTGAATGATGAAGAAAAACCCCTGTACATTACATGAAGGC 2048
681 sProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlyS 698
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2049 AGACCGGCTATTGGTCCATTGCTCCCAATGCTCTGCTTTGGCTGCCT... 2094
698 etArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGly 714
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2095TGCCCTCGATGGACCGGATCGAGTGGAGATCCA 2127
715 ArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla.....Al 728
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2128 GCGGACGACAGTGTGAAGGATAGGAGGGAGCTGGAGAGCAGTCACGC 2177
728 aGlyTrpLeuThrAlaLysHisSerAspGlySerLeuSer 742
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2178 AGGAGCCCTCGCGTGATTTCATTCATGAACAAGCTGGAGTCAGT 2220

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:US-09-508-377-13

seq_documentation_block:
; Sequence 13, Application US/09508377
; GENERAL INFORMATION:
; APPLICANT: KALEEN, ZHONGYILI
; APPLICANT: MORELL, MATTHEW
; APPLICANT: RAHMAN, SADEOUR
; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS

```

; FILE REFERENCE: 054270/0126
; CURRENT APPLICATION NUMBER: US/09/508,377
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: AU PP 2509
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/AU98/00743
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: AU PP 9108
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 10337
; TYPE: DNA
; ORGANISM: Triticum tauschii
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (10232)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-508-377-13

alignment_scores:
    Quality: 2346.00      Length: 3317
    Ratio: 3.141          Gaps: 25
    Percent Similarity: 22.520      Percent Identity: 22.339

alignment_block:
US-09-674-824-2 x US-09-508-377-13 ..

Align seg 1/1 to: US-09-508-377-13 from: 1 to: 10337

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
|||||
1 ATGGCGGCGACGGCGGTGCGGCGCGGGTGCTCGCCCCCAGCGTCCGCT 50
|||||
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
51 GCAGCGCGATCCGCGACGCGCGCGCGGGGTCCGCTTGGCTTCCGCGC 100
|||||
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
101 CGCGGCTCCCGCGGCTTGGCGGGGCGCGCTACGTGCGGAGCTCAGCAGG 150
|||||
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
151 GAGGCGCCCGCGGCGCGCGCGCGCGCGCAACTGGCGCGCGCGCT 200
|||||
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
|||||
201 CGTGCCAGGTTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 250
|||||
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
251 CCGCGAGCAGCGCGCGCTGCCGAGCGCGCGGTGGGGGAACCTCGCGCC 300
|||||
101 AspLeuLeuLeu.Glu..... 105
|||||:::
301 GACCTCTGTCGAAGGTAAAAAACAAGGCTGAATCCTCAGATCACTCCG 350
|||||
105 ..... 105
351 CGTCTTCGTTTTACAAATACGGTACTCGAAGTGGTCTGTATATGTGA 400
|||||
105 ..... 105
401 AGTTTCTCGATTCTTCCTGACGGATGTTTCAGTCGATTCAAGTTGTATA 450
|||||
105 ..... 105
451 TATGTGATACGTTGTTGTTGTTTCATCGATCGTACAGATTTTACCAGCACACTA 500
|||||
105 ..... 105

```

1401 GTCTGCTAATATCTGTCTCTGCTGCTGTGACAAATATACCATATGA 1450
106 GlyIleAlaGluaspSerIleAspSerIle 115
1451 ATATTTTCCATTTTCCACACCGGATGCTGAGGATTCCTCGACAGCAT 1500
115 eileValAlaAlaSerGluInAspSerGluIleMetAspAlaAsnGluG 132
1501 AATCTGGCTGCAAGTGAGCAGGATCTGAGATCATGATGCGAATGAGC 1550
132 InProGlnAlaLysValThrArgSerIleValPheValThrGlyGluAla 148
1551 AACCTCAAGCTAAAGTTACAGCTAGCATCGTGTGTTGACGTGGTGAAGCT 1600
149 AlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuPr 165
1601 GCTCTTATGCAAGTCAAGGGGCTGGGAGATGTTGTGGTTCGTACC 1650
165 orleAlaLeuAlaAlaArgGlyHisArgValMetValValMetProArgT 182
1651 AATTGCTCTTCTGCTGCTGCTGCTACCGTGTGATGTTGTAATGCCAAGAT 1700
182 TyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAla 198
1701 ACTTGAATGGTCTCTGATAAACACTATGCAAGGCATATATACACTGCG 1750
199 LysHisIleLysIleProCysPheGlyGlySerHisGluValThrPhePh 215
1751 AAGCACATTAAGATTCCATGCTTTGGGGATCACATCAAGTGACCTTTT 1800
215 eHisLluTyrArgAspAsnValAspTrpVal...phe..... 226
1801 TCATGAGTATAGAGACAACGTCGATTGGGTGGGTACACAAATCACTTCTT 1850
226 226
1851 ATTCTCTGTTGAATGTAGCACTGTTATATCCCTGTTTACACTCTTTTA 1900
226 226
1901 GCCCTGCAAGACATATGTGATTTCACATACTTTTGTATTATTCCTTGT 1950
226 226
1951 ACTCTTGCTCATGAGGTCAAAATATCATATATCCATGGAAGTCATGCAT 2000
226 226
2001 GTCCCTAGTATTTTGTGTGCTGCTGCTTTAACTTTCAGGGATTAATACG 2050
226 226
2051 TGGAAATTGTAATACTAAAGTTTATTTTATTTGAAAAAATTTGAGTTGGC 2100
226 226
2101 TGAGCCACAGCAGCAGTCAGTGGCACCACCTGCTTGCACATGATTTGCATT 2150
226 226
2151 TCTGTTGACCGAGCACCTTCATGTGAATGAAGTGTAAATATCAATAAGTA 2200
226 226
2201 CCAATTTTATCTGCCAATTGCACITTAAGAGTATATACATTTATCTTGGC 2250
226 226
2251 CTCAATCATGGGAGTACTGTGCATTCAAGTGCACCATCATTTGTTCTAAGGA 2300
226 226
2301 GAAATGTGGGTGCAAGGAAGACACTTTTGTCCCTTAATAAAGGAGCGC 2350

226 226
2351 ACTCTGTTGTCAATATAGATGAAGACAACAACTTATTTCAAGAGCTAA 2400
226 226
2401 CAATGGCAAAAGAACCAAAAGAGCATGTAAAGCGGTGACACCAAAAGG 2450
226 226
2451 TGAGGGGGGCTTGTGACTCACAGCACCCCAAACTATTGCCATTTGTTTA 2500
226 226
2501 CTAATGAAGATCATTTTAGAAGCTCTCAGGAACCTTCGAAAAACAGTGGCT 2550
226 226
2551 TTCGGTCCACAGATCGTCTGCTTAAATATTTTGTCCAGTGATACTTTTTT 2600
226 226
2601 GCTCCTTACAAGAGTGCCTATGTTGACATATACATTTGTTAAGTTGTTTCA 2650
226 226
2651 AAGTTTACTTCTTATTCTAAACAGCAAGTCCCTAAATGCTTGCAATTTATT 2700
226 226
2701 TGCGTATTTATTTTATTTCTCATTTTCAATCAACACTTTTGTTCAGGTGT 2750
227 ValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsn 243
2751 TGTCTGATCATCGTCTCATATCATAGACCAGGAAGTTTATATGGAGATAAT 2800
243 heGlyAlaPheGlyAspAsnGln...phe..... 251
2801 TTGGTCTTTTGGTGATAATCAGGTACACTACACTATATACTAAGCTCTAG 2850
251 251
2851 TTGACTAAGTCGTAAGTTGTACCTCTCGCTGACCGGCTGCTCTATGTCG 2900
252 ArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeu 265
2901 TGCAGTTTCAGATACACACTCTCTTGTCTATGCTGTCATGGAGGCCCACTA 2950
266 IleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValVa 282
2951 ATCCCTTGAATTGGGAGGATATATTATATGACAGAATTCATGTTTGTGT 3000
282 IasnAspTrpHisAlaSerLeuValPro...Val..... 292
3001 GAACGATTGGCATGCGACGCTTGTGCCAGTGTACGTTGTTGTGGATCTG 3050
292 292
3051 AAAGTCCCAATCCTTTATTCATTTCTGCTTTTGCAGTGTGCCCATGTCTAC 3100
292 292
3101 ATTTCTTTTATGCTTTTTCATGCTGTTCTTATATATTCATATATGCTTA 3150
292 292
3151 TGGAGTCTAAAAAGTTACGGAGGAATAACTCTTAAGGATTTTCCTCAATC 3200
292 292
3201 AATTATCTTTAGCTTTAGTTAACATTTTACTGTGGCAACATAAATGTGTTT 3250

292 292
3251 TGAGATTACAAGTTTCAGAGATTCACACTTCACTAGTTCGTAGCTAATCTG 3300
292 292
3301 ATGTTTTCCCGAGAGAAATGCCTAAAGCTTTGTGTCTTGATGCATTGATA 3350
292 292
3351 GAAAAAGAGTTTATGTACACTCCCAAGAGGGGACCCAAATTTACACAC 3400
292 292
3401 CACACCCTGAGAACTAGGCGCTGCCGGAAGACCGATGCAAGCCCACT 3450
292 292
3451 GCCCCTGCCTTAGCTCAAAGCGGCGTCAGCTTGATTGTCAAGTAAG 3500
292 292
3501 CTAGCAGTGTAGATTGCGCAAGTGCATTCGTGGAAGATGACAGTGTG 3550
292 292
3551 CGCTGCTTCCAAATCCACCNAACTATGAGCATGATCACTGGAGAAGTACC 3600
292 292
3601 TTTTCTCGCGCTGAGGGGTGGACTGGTGTCTGCTGCCAGTTTTC 3650
292 292
3651 AGATAAATCTGAAAAATGCATGTTTGTGATGATTTAGTATCTTGGCGACC 3700
292 292
3701 TGGGTACCACCTAAGCTTTCACACAGTAATTTGACAGTTACACCTATAAAA 3750
292 292
3751 GTAAGGTCATGATATGCATGTGTTTGGGTAGATCATGGTGCATGCATT 3800
292 292
3801 TTAGGAATTAGGACATGCCAGAACCGTGAGGCTTATGGGCAATTTCAT 3850
292 292
3851 TTGTTCCATTATACGAGTCATGAATATGGTTACAGCATGTTTGGACGTAC 3900
292 292
3901 TTGTTTGGGCAATTTACAGTGGTGAATTTGAGTGCTTGATGTTGGCTA 3950
292 292
3951 GCTGGCTTATTTGTACAAGTATCGATGTAGATGCATATTTCCCTTTTGT 4000
292 292
4001 TCTTGTGCTGTTTGGCATGTTGTATTTCCCTTTTCTGCGCAGTGTTCG 4050
292 292
4051 ATGTTAAATTGCTTTTCATTACATAATCAACTTTGTTGCTGACATCAGTC 4100
293
4101 ATTTTATTCAGCCTTCTTGTGCTGCAAAATATAGACCATACGGTGTTCACA 4150
305 rgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGln..... 319

4151 GAGATTCCCGCAGCACCTTGTATTACATAATTTAGCACATCAGGTTTGG 4200
319 319
4201 GTCATACACCTTTCATTATCCGTACATGGCTTTGTAAAGTCGGTTCACACG 4250
320 Gly.ValGluProAla 324
4251 TATCGTCATACTGTATGTTATTTCAATGTTCATTAGGTTGTGGAGCCTGCA 4300
325 SerThrTyrProAspLeuGlyLeuProGluTrpTyrGlyAlaLeuGln 341
4301 AGTACATATCCTGATCTGGGATTGCCCTGAATGGTATGGAGCTTTAGA 4350
341 utrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyGlu 358
4351 ATGGGTATTTCCAGATGGCAAGGAGGCATGCCCTTTGACAAGGCTGAGG 4400
358 laValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThr 374
4401 CAGTTAACTTTTGAAGGAGCAGTTGTGACACACAGATCGAATTTGTGACC 4450
375 ValSerGln..... 377
4451 GTCAGTCAGGTGAATACTCAATCTCTCTTTTTCITTTGCGGGATGTT 4500
377 377
4501 CTTTCAGTTCAATTCGCCCTGTCTTTTACCCTAATTAAGAAATGATTTAATCT 4550
378 GlyTyrSerTrpGluValThrAlaGluGlyGln 390
4551 TTTGTTTCTAGGGTTATTCATGGGAGGTCACAACTGCTGAAGTGGACAG 4600
391 GlyLeuAsnGluLeuSerSerArgLysSerValLeuAsnGly Ile. 406
4601 GGCTCAATGAGCTCTTAAGCTCCCGAAAAGTGTATTGAATGGTAACATA 4650
406 406
4651 TATTTGAATCCACTTATCTTCTTGAACATATTTACAGAAATAGATGG 4700
407 Va 407
4701 ATGGGTTGCAAGAATAAATTCAGTTTGCCTCTTTCGGTATCAAGGAATGT 4750
407 IasnGlyIleAspIleAsnAspTrpAsnProThrAspLysCysLeuP 424
4751 AAATGGAATTGACATTAATGATTGGAACCCACACACAGACAAGTGTCTCC 4800
424 roHisHisTyrSerValAspAspLeuSerGlyLys..... 435
4801 CTCATCATTTATCTGTCGATGACCTCTCTGGAAGGTGTGTGATAGTAC 4850
435 435
4851 CCTATATAAACAATGATATATCTGATCTAGTACTTTCTTTTCTTTGCTA 4900
435 435
4901 GTTTGCTTCCATGATGTTCTCACTAACTAATCTATCTGTGTTTGGCATA 4950
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuPro 449
4951 CTTGTCAGGCCAAATGTAAAGCTGAATTTGCAGAAGGAGCTGGGTTTACCT 5000
450 ValArgGluAspValProLeu..... 456
5001 GTAAGGAGGATGTTCTCTCTGTTAGATACAAACCCCTTAAGATATATATT 5050
456 456

5051 TTTTAAATCCCTAAAAAAACTTGGCGATCATCTCATTTAGCTTGATTCAC 5100
457 ..IleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuile 472
|||||
5101 AGATTGGCTTTATTGGAAGACTGGATTACCAGAAAGGCATTGATCTCAT 5150
473 LysMetAlaIleProGluLeuMetArgGluAspValGlnPhe..... 486
|||||
5151 AAAATGGCCATTCCAGACTCATGAGGGAGGACGTGCAGTTTGTAAGTTC 5200
486 486
5201 ATATTCTTTTCTTGAGACTAGAGTATATAATCAAAACATGTAGGTGGGG 5250
486 486
5251 TGGTATATACAGACATAAGTTCCAGCTATTGCTTCCATGAGAATTTAA 5300
486 486
5301 TGCATTTCAGTAATATGCTACTGCAAGTTTGTAAACAAAGTTGGAAGCAA 5350
486 486
5351 TAAATATATGTAGCAGTACCATGACATGCAGTGCCTACTATAGCTGAATGTC 5400
486 486
5401 CTGTAGTCTATGTATCAACACTCAACACATGTTTCGCATACAAA 5450
486 486
5451 CACATGCGTGGCGCAACAACATACTCTACAATAAAATTTGGCTTGCTGA 5500
486 486
5501 ACTGCAGACATGCTCTTATCTCTCCATTTCCAACATTTCTTGTTCACATTC 5550
486 486
5551 GCTAAGACTAAGAGAGGGGGAGCCAGGGTGATGTAGCCAACATAGATCC 5600
486 486
5601 AGTAAGGAAGCTAGCCGAGCCTAGGAGGATTCGCTTAGGTAGCTGGAACG 5650
486 486
5651 TAGGGTCTCTGACAGGAAGCTTCGGGAGCTAGTCGATGCAGTGGTGAGG 5700
486 486
5701 AGAGGTGTTGATATCCTTTTCGCTCCAAGAAACCAATGTAGGGACAGAA 5750
486 486
5751 GGCGAAGGAGGTGAGGATACCGGCTTCAAGCTGTGTTACATGGGACGGC 5800
486 486
5801 TGCAACAGAAATGGCGTAGGCATCTTGATCAACAGAGCCCTTAAGTATG 5850
486 486
5851 GAGTGTAGACGTCAAGAGACGTGGGGACCGGATTATCCTCGTCAAGCTG 5900
486 486
5901 GTAGTTGGGACTTAGTTCTCAATGTTATCAGCGTGTATGCCCCCAAGT 5950
486 486
5951 AGGCCACATGAGAACGCCAAGAGGGAGTTCTGGGAAGGCCTGGAAGACA 6000

486 486
6001 TGGTTAGGAGTGTACCGATTGGCGAGAAGCTCTTCATAGGAGGAGACCTC 6050
486 486
6051 AATGGCCACGTGGGTACATCTAACATAGGTTTGAAGGGGCACATGGGG 6100
486 486
6101 CTTTGGCTATGGCATCAAGATCAAGAAAGAGATGTCTTACGCTTTGCTC 6150
486 486
6151 TAGCCTACGACATGATTGTAGCTAACACCCCTCTTTAGAAAGAGAGATCA 6200
486 486
6201 CATCTGCTGACTTTTAGTAGTGCCCAACACTAGCCAGATCGATTTTCATCC 6250
486 486
6251 TCTCGAAGAGAGAGATAGGTGTGCGCGCTAGACTGCAAGGTGATACCT 6300
486 486
6301 TCGGATTTCGTCCAGCGGATAAGCGTGCCAAAGTCGCTAGAAATGAAGT 6350
486 486
6351 GGTGGAAGCTCAAGGGGAGGTAGCTCAGGCGTTCAAGGAGAGGTTCATT 6400
486 486
6401 AGGGAGGGCCCTTGGGAGGAGGGGATCGGACAATGTGTGGATGAA 6450
486 486
6451 GATGGCGACTTCATTCGTAAGTGGCCCTCGGAGGAGTGTGGAGTGTCCA 6500
486 486
6501 GGGGATGGAGAAGCGAAGATAGGATACCTGCTGGTGGTAATGATGTC 6550
486 486
6551 CAGAAGGCAATTAAGAGAAAGAAAGATTGCTTTAGACGCTATACTTGA 6600
486 486
6601 TAGGAGTGCAGTCAACATAGAAAACTACAAGATGGCGAAGAGGCCGCAA 6650
486 486
6651 AGCGAGCTGTCAAGTCAAGGAGGGTGGGCATATGAGGATCTCTACCAA 6700
486 486
6701 CGGTTAGCAGCAAGGAAGGCGAAAGGACATCTATAAGATGGCCAAGAT 6750
486 486
6751 CCGAGAGAGGAAGACCGAGGATATTGGCCAAGTCAAAATGCATCAAGGA 6800
486 486
6801 TGGAGCAGACCAACTCTTGGTGAAGGACGAGGAGATTAAAGCATAGATGC 6850
486 486
6851 GGGAGTACTTCGACAAGCTGTTCATGGGGAGGATGAGAGTCTTACCATT 6900


```
8701 TATCAGCTTGGATATATATATAATGTTCAAAACATTTATGTCCTCTTTT 8750
529 .....CysAspIleLeuLeuMetProSerArgPheGluProCysGlyL 543
8751 GTGCAGTTGCGATATATGTTATGCCATCCAGGTTTGAACCTTGTGGTC 8800
543 euAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGly 559
8801 TTAATCAGCTATATGCTATGCAATATGTACAGTTCTCTGTAGTTACATGA 8850
560 ThrGlyGlyLeuArg..... 564
8851 ACTGGGGCCCTCCGAGTAAGACAACCTGCTTGAATAATTTATGTTATCTTG 8900
565 .....AspThrVal 567
8901 GCTCCAACGCAAAATGTTCTAATTTGGCTCGTGTATTCAACAGGACACAGTC 8950
568 GluThrPheAsnProPheGlyAlaLysGlyGluGlyThr.Gly.... 582
8951 GAGACCTTCAACCCCTTTTGGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 9000
582 ..... 582
9001 ACTGCTCAATTTAGCTAACTTTTTCAGTTTATCTTTTGGCAATGCTTGGG 9050
582 ..... 582
9051 GGTTCATTGCGCCATAAATCAACTTGTGATAATTAACCTGTTACTGTTCTG 9100
583 .....TrpAlaPheSerProLeuThrValAspLysMetLeuTrp. 595
9101 TACTTGAGGTGGGCGTCTCACCGCTAACCGTGGACAAGATGTTGTGGG 9150
595 ..... 595
9151 TAACTTTTGTGAGCTCTTGTCCGGTTATAGGATCGACCTTGCTGTAG 9200
595 ..... 595
9201 CATGGTACCTTAGTGCCCTTGTATATAGACCTAACCTGTAGGACTCACT 9250
595 ..... 595
9251 TTGCTACACTAATCATAGTAGTCGATTGCCGGAGCGGTTTGTCTGGA 9300
595 ..... 595
9301 TTCTGCTAATTTAATTTTCATGACATACTCATACCATGTTTGTGTTCT 9350
595 ..... 595
9351 CCGATGGGCGCCAGAAATGGCGTCTAGTGTCTGCGATCTGTGTAACTAGCC 9400
596 .....Ala. 596
9401 AATGCGCGGTTGTTCCAAGTGAATAATTTACCTTTTGACCAATTTGTCAGGC 9450
597 LeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGluG 613
9451 ATTGCGAACCGCGATGTCGACATTACAGGAGCACAAGCCGCTCTCTGGGAGG 9500
613 lLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla.Pr 629
9501 GGCTCATGAACGAGGCGATGACGAAGAACCATACGTTGGAGCATCCGCC 9550
629 oSerSerThrArgSerSerGlyProSerTrpThrAsnProThrS 646
9551 GAGCAGTACGACGATCTTCGAATGGGCTTCGTGGAGCAACACCTACCT 9600
646 erCysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaLeu 662
9601 CATGTAGACGGGACTGGGGAGGTCGAAGCGGGGTCTCTCTTGAGCTCTG 9650
```

```
663 LysThrSerSerSerPheArgGlyProGluGlyTyrProCysThrLe 679
9651 AAGACATGTTCTCTCATCTCCGCGGCCCGAAGGATACCCCTGTACATT 9700
679 uArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheA 696
9701 CGGTGTGCTCTGTACAGTAGAGTCGAATGCGCCTGCTTGTCT.TGGTCCG 9749
696 lagLysSerArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSer 712
9750 CCGGTTGAGAGTAGATGACGGCTGTGCTGCTGCGGGGTCGACAGCTCG 9799
713 GlyGlyArgGlnLeuGlnPheTrpGlyLeuArgLysGlyCysAlaAlaG 729
9800 GGTGATGACAGATTACAGTTTGGGGAATAAGGAAGGATGTGCTGCAGG 9849
729 yTrpLeuThrAlaLysHisHisSerAspGlySerLeuSerValArgVal 746
9850 ATGGTTAACAGCAAGCACCACTCAGATGCGAGCCTCTCTGTCGCTGTTA 9899
746 hrAlaGluIleArgAsnGlnLeuValThrLeu 756
9900 CAGCTGAATCAGAAACCAACTGGTGACTCTT 9931
```

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq1:US-09-815-264-86353

```
seq_documentation_block:
; Sequence 86353, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 86353
; LENGTH: 16095
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(16095)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-86353
```

alignment_scores:
Quality: 1523.00 Length: 1887
Ratio: 3.058 Gaps: 27
Percent Similarity: 26.391 Percent Identity: 23.953

alignment_block:

US-09-674-824-2 x US-09-815-264-86353/rev ..

Align seg 1/1 to reverse of: US-09-815-264-86353 from: 1 to: 16095

```
159 AspValCysGlySerLeuProfileAlaAlaArgGlyHisArgVa 175.
16093 GATGTTTGGTTCACTCCCAATTCCTCTGCTCTGTTGGTCATCGTGT 16044
175 lMetValValMetProArgTrpLeuAsnGlySerSerAspLysAsnTyrA 192
```

|||||GATGGTGTAAATCGCGAGATACATGAACGGGCCCTTGACCAAAATTTTG 15994
16043
192 laLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGly 208
192
15993 CAACAGCATTTTACACTGAGAGACACATTAAGATTCCATGCTTTGGCGGA 15944
15993
209 SerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrp.. 224
209
15943 GAACATGAAGTTACTTTTTTTCACGAGTATAGGATTTCTGTTGATGGGT 15894
15943
224 224
224
15893 GAGTATTCATCTCCTTACTATTTTCTGATGAACAAACAACATATCTGCTAT 15844
15844
224 224
224
15843 CTTTTTGTCTTATCCCACTGAAGACATGATCATATATCTGTAAT 15794
15843
224 224
224
15793 CTGTTTCTATGCTTACTGCTATACACTTGTGCACCTGGTTGATGCTGG 15744
15744
224 224
224
15743 TGTTCAGATCAAAATGCAAAAGATTCCCTCATATGACTTAAATGTATCC 15694
15743
224 224
224
15693 ATTTTATGCTCAAGCATTTGTGACTTCTGTTTGCACAAGCATTAA 15644
15644
224 224
224
15643 TTAATAAGTTGTACTGTTGGCTTTCCATTGCCATTAAATGATCAGTA 15594
15643
224 224
224
15593 ACATTTAGTTATATATCATGGGAATTATGTACTTCAATGCCATTTTATTT 15544
15544
224 224
224
15543 CTTGAAAGTGAACATGTGGTTTACATTGAAACATGTATGTGCACAAC 15494
15494
224 224
224
15493 TAATAGATGTAGATGGCATGACATGATATATCAATTAAGGAACAAGATGTG 15444
15444
224 224
224
15443 TATAACCTAGGAGGAACATTTCTAGCATTTTGGTTGCAAGATTGATTC 15394
15443
224 224
224
15393 CTTGTTGAGTTTACCAATGTTTACTTATTCCTATTTGGAACTTGGTAGTG 15344
15393
224 224
224
15343 TTCACCTCCGCGACCATCACCCATGACACACATGCTGATTGTAGAAAG 15294
15343
224 224
224
15293 CATACCTGCAATTTGGGTACTTATTTTGTGTACACCTCATTCTCCCTG 15244
15293
225ValPheValAspHisProSerTyrHisArgProGlySerLeuTyr 239
225
15243 TTCAGGTGTTGTTGATCATCCCTCATATCATAGACCTGGAATTTGTAT 15194
15243
240 GlyAspAsnPheGlyAlaPheGlyAspAsn..... 249
240
15193 GGAGATAAATTTGGTGTCTTTTGGCGATAATTCAGGTATTTCTTACATCAAGA 15144
15193
249 249

15143 TCCTATTTGACCAGTCTAAATATATATCACAATGAGTCATTGATTGCACATT 15094
15143
250GlnPheArgTyrThrLeuLeuCysTyrAlaAla 260
250
15093 GACTTCCCTATATATCTGAGTTCAGATACACACTCTCTGTGCTATGCGGG 15044
15093
261 CysGluAlaProLeuIleLeuGluLeuGlyTyrIleTyrGlyGlnAs 277
261
15043 TGTGAAGCCCATTAATTTCTTGAACCTGGGAGGATATATCTATGGACAGAA 14994
15043
277 nCysMetPheValValAsnAspTrpHisAlaSerLeuValProVal..Leu 293
277
14993 ATGATGCTTTGTTGTGAATGATTGGCATGCCAGCTTTGTGCCAGTGTATG 14944
14993
293 293
293
14943 TTCTGGATCTGAACCTCTGCAACCAAAATTGATTAAACATTGAATATCTCTT 14894
14943
293 293
293
14893 ATAAACCATTTTTTTCATACGTTTGTCTTTGTGTTGTGTGTCTATGTT 14844
14893
293 293
293
14843 AGATAAGAGGAACAAATATATTAGATTGAGTGAATAGGAGGATTTTCTCAA 14794
14843
293 293
293
14793 TCATTTTTTTTCTCAAAAGCGCAGGAGAACTGGGATTTCATTTTCATTATTT 14744
14793
293 293
293
14743 TCTCAATCATTTTGATATTTCCAAAAACATGAACAGTCCAGTACATCATTT 14694
14743
293 293
293
14693 TATGTGTTTAAATGATTTAGTTTGTGCTACACTAGTTTCATAGCTGATCT 14644
14693
293 293
293
14643 GAAATCTCCCTGCCTATGTTTTTTTTTGGAAATATTTCTGAGTAGTGC GA 14594
14643
293 293
293
14593 AGTTTCTTATTTATGGTAGTGCCTATTTTGTAGTTTCATCTTCTACAGATACA 14544
14593
293 293
293
14543 ATTCAATACCCAAAGATTTGGGAAAAATCCTGTAGTCTAAATGGCCCTAT 14494
14543
293 293
293
14493 CATGCTACTAGTGACTGAGGTGATAGTTTCCAAACCCCTCTAATTTGTGGT 14444
14493
293 293
293
14443 AAATTAAACTGCTGTTTTCACACCATAATTTGATGTTTATAATTATTAAAG 14394
14443
293 293
293
14393 AGAGTGGTGGTACATGCTTTTCTAACTTCTAAGACACTTCAGAACACAT 14344
14393
293 293
293
14343 GAGGTTAATGGTGAATGCTTTCCCTGCATATATTGAATATGGTTTGCAT 14294
14343
293 293
293
14293 CCATCAGATGCTACATTTGCTACTCCTCTTTTGTGGCATGTTCTCCACGGT 14244
14293
293 293
293
14243 GGATAAGATACTAATGCGCAGTGAAAAACTCATCACACTTATTTTTTTAGTCT 14194
14243

293 293
14193 CCTTTTACTCAATGCTGAATGTTAAATTTGGTTTTCCTTACATATCAAT 14144
294 LeuAlaAlaLysTyrA 299
14143 TTAATATAACTTTTATCTGTTATCCAGCCTTCTTCTGCAAAATATA 14094
299 rGProTyrGlyValTyrArgAspSerArgSerThrLeuValIleHisasn 315
14093 GACCATATGGTGTACAGGATGCCGCCAGTGTCTTGTCATACATAAT 14044
316 LeuAlaHisGln..... 319
14043 CTAGCACATCAGGTTTGAGCCTATCATTTTGCATATCTTTTAACTAGTTT 13994
319 319
13993 TGTAAATGCCACTTGTAAAGTCCAGTATGTAGTATTTTCATTTGTGTTCAA 13944
320 GlyValGluProAlaSerThrTyrProAspLeuGlyLeupro 333
13943 TGTATAGGTTGTGGAGCCTGCCAGTACATATCTGACCTGGGATGCCA 13894
334 ProGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgAr 350
13893 CCTGAATGGTATGGAGCATTAGAATGGTGTTCACAGTGGGCAAGCGC 13844
350 qHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValV 367
13843 GCATGCCCTTGACAAGGGTGAGGCAGTCAATTTTTTAAAGGGCGAGTTG 13794
367 alThrAlaAspArgIleValThrValSerGln..... 377
13793 TGACAGCAGATCGAATTTGACTGTCCAGCCAGGTGAGATTCCTTCCTTTT 13744
377 377
13743 CGTCTACTAGTTGGCTTACGGCCACAGTCAAAATTTGAAATTTGAACCTA 13694
377 377
13693 ACCTTAGAGTTGATTTTTTTTCCCATCCTAGTCTATTTTATGCCCTGG 13644
377 377
13643 CTTTAAACCCGAAATTTCTTCTTCATTTTGTCTTGCACTCAATTAAGGA 13594
378 Gly.TyrSerTrpGluValThrAlaG 387
13593 ATCATTTATGTTTCATACCAGGGGTATTCATGGGAGGTCACAACTGCTG 13544
387 luGlyGlyGlnGlyLeuAsnGluLeuSerSerArgLysSerValLeu 403
13543 AAGGTGGCAAGGCCCTCAATGAGCTCTTAAGTCCCGGAAGAGTGTATTG 13494
404 Asn..... 404
13493 AATGGTACCTATATTTGGGTTTTTATATTGTTGTTCTTGAGCTGATCATTG 13444
405G 405
13443 TAATAGGTGGATGATACAAATTAATCAATCTATTTCCTGGATATGTAG 13394
405 lyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLys 421
13393 GAATGTAATGGAATTGACATTAATGATTGGAACCCATCCACAGACAAG 13344
422 CysLeuProHisHisTyrSerValAspLeuSerGlyLys..... 435
13343 TTTCCTCCCTTATCATTTCTGTTGATGACCTGTCCGGAAGGTGAAGTA 13294

435 435
13293 TAGTACCATGATATACAGATTGTCATCTCGTTTCATTTTTTCTTTGCA 13244
435 435
13243 CCTCGTCGTTTCTTTCACAGACTAATTTATGAAACCTCGAATTCATGTTT 13194
436 Ala..LysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArg 451
13193 AGGCCAAGTGTAAAGCTGAATTCAGAAGAGGCTGGGTTTACCTATAAGG 13144
452 GluAspValProLeu..... 456
13143 CCCGATGTGCTCTGCTGTAGATACAAATATTTATGAAATTTATTTAT 13094
457 IleGlyP 459
13093 GTATTCTTTTATGGGCAGTTGGATTATTAGCTTTATTTACAGATTGGCT 13044
459 heIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAla 475
13043 TTATTTGGAGATTGGACTATCAAAAAGGCATTGATCTAATTAACCTGCC 12994
476 IleProGluLeuMetArgGluAspValGlnPheVal..Met..... 488
12993 ATTCCAGATCTCATGGGCAATATTTCAATTCGTAAGTTCTGTTTCTTT 12944
488 488
12943 TTTCTGCTGAGACTAGAAAGTCAATGTGATGATAGTCTTAGTGTATAGT 12894
488 488
12893 AGTATATTGACACTTATATAGCTTGTATCATGAATCTAAGGAATGTAAT 12844
488 488
12843 ATGCAGCTATTGCTTTTCATTGTCAACGCATATGCAAGTAAACTAGTATT 12794
488 488
12793 CAGATCAAAATATTAGTGTCTATAAGAAAAATGCATATGCAAGTCATAA 12744
488 488
12743 AGAAGATAGCAGTAGTATAAATTTGACCACCTGTATGACTTATTAGCAACC 12694
488 488
12693 ATCCAGTGTGATATCAATTTGGAGCACCTGCTTGGCCAGCCAGTCTGTTAT 12644
488 488
12643 TTCCACCTTTTATACATTTNGGTACACTTAAACAAACATCTTCTGACAGAC 12594
488 488
12593 ACCAGGGAGCGGTATACATGCAGACAAAAGGATCACCATTATTATTAAGTC 12544
488 488
12543 AGTAGGATCTAACTTTAGTAGCAGCAAAATGCTTTGTATGAATGCAACCACG 12494
488 488
12493 TGCCTTATCCCAATACTTCATTTGTTGGGGGTTTCTGGCTTTTCAACTGGAG 12444
488 488
12443 ACTAAACAGAGTAGTGTGTTGAGTGGGTTTAAATAATTTTATCCACATAT 12394
488 488


```
1727 GGCCTGAAGATGCGGACAGGTTGTGCTGAGCCCGGGTACCTGTG 1776
381 pGluValThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSers 398
1777 GGAGCTGAAGACGTGGAGCGGCTGGGGCTTCACGACATACATACGGC 1826
398 eArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAsp 414
1827 AGAAGCACTGGAAGACCGCGGCATCGTGAACGGCATCGACATGAG 1876
415 TrpAsnProThrThrAspLysCysLeuProHis.....HisTy 427
1877 TGAACCCGAGGTGGACGTCACCTCAAGTCGACGGCTACACCAACTT 1926
427 rSerValAspLeu...SerGlyLysAlaLysCysLysAlaGluLeuG 443
1927 CTCCCTGGGGAGCTGGAGCTCGGCAACGGCAGTGCAGGAGGCCCTGC 1976
443 lnLysGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPhe 459
1977 AGCGGAGCTGGCCCTGCAGGTCCGCGCGACGTGCTCGCTGCTGCTTC 2026
460 IleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaI 476
2027 ATCGGGCGCTGGACGGGCAAGAGGGCTGGAGATCATCGCGGACGGAT 2076
476 eProGluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyA 493
2077 GCCCTGGATCGTGACCGAGGAGCTGCACCTGCTGCTGGCCAGCGGCG 2126
493 spProfilePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAsp 509
2127 GCCACGACCTGGAGGAGCTGCTGCGGCACTTCGAGCGGGAGCACACGAC 2176
510 LysPheArgGlyTrpValGlyPheSerValProValSerHisArgIleTh 526
2177 AAGGTGCGGGGTGGGTGGGTTCCTCGGCTGGCGTGGCGACCGGATCAC 2226
526 rAlaGlyCysAspIleLeuMetProSerArgPheGluProCysGlyL 543
2227 GCGCGCGCGGACGGCTCTCATGCCCTCCGGTTCGAGCGCTCGGAC 2276
543 euAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValHisGly 559
2277 TGAACCACTCTACGCCATGGCTACGGCACCGTCCCGCTCGTGCATGCC 2326
560 ThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyAla 576
2327 GTCGGTGGCTGAGGAGACCGCTGGCGGCTTCGACCCCTTC..... 2368
576 sGlyGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysM 593
2369 .AACCACTCGGCTCGGTGGAGTTCGACCGCGAGAGCGCAGAACG 2417
593 etLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysPro 609
2418 TGATCGAGGCGCTCGGGCACTGCTCCGACCTACCGGACTACAGGAG 2467
610 SerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAs 626
2468 AGCTGGAGGGGCTCCAGAGCGGCGATGTCGAGGACTTCAGCTGGGA 2517
626 pHisAla.ProSerSerThrSerArgSerSerSerGlyProSerTrpThr 642
2518 GCATCGCGCAAGCTCTACGAGGACGCTCCTCGTCAAGGCCAAGTACCA 2567
643 Asn.....ProThrSerCysArgArgG 650
2568 GGTGAACGCTAGCTGATCGCGTCCAGCCCGCGCATGCGTGCATGACAG 2617
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
2618 ATCGAAT.....TCCGCAATTCGCACGACGAGGAAG 2646
```

```
667 erSerPhe.....ArgGlyProGluGly..... 674
2647 GTCCATGGAGCGCGCGCATCCCGAAGTACAGTGACATGAGGTGTGTGT 2696
674 ..... 674
2697 GGTGAGACGCTGATTCCTGATCTGGTCCGTAGCAGAGTAGCGGAGGTA 2746
675 .....TyrProc 677
2747 GGGNAGCGCTCTGTGTACAGGTATATGGGAATGTTTAACTTGGTATT 2796
677 ysThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeu 693
2797 GTAATTTGTTATGT.....TGTGTGCAT 2819
694 TrpPheAlaGlySerArgThr.....TyrAspGlyCysAlaAl 706
2820 TATTACAGAGGGCAACGATCTGCGCGGCGCACCGGCCCAACTGTTGGGC 2869
706 aAlaAlaValThrAlaSerGlyArgGlnLeuGlnPheTrp 720
2870 CGGTGCGACAGCAGCGCTGGATCCGACCGCTGGGCGCTGG 2912
seq_name: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-044-543-25
seq_documentation_block:
; Sequence 25, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthesis Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches
; FILE REFERENCE: 1144D
; CURRENT APPLICATION NUMBER: US/10/044,543
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 2418
; TYPE: DNA
; ORGANISM: Typha latifolia
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2418)
; US-10-044-543-25
```

```
alignment_scores:
Quality: 1201.50      Length: 643
Ratio: 2.814          Gaps: 15
Percent Similarity: 66.407 Percent Identity: 42.457

alignment_block:
US-09-674-824-2 x US-10-044-543-25
..
Align seg 1/1 to: US-10-044-543-25 from: 1 to: 2418

42 GlyArgTyrValAlaGluLeuSerArgGluGlyProAlaAlaArgProAl 58
|||||:|||||:
325 GGGAAGTACAAAGC.....GGCGCGTTCCTAA 353
58 aGlnGlnGlnGlnLeuAlaPro.....ProLeuValP 69
|||:|||||:
354 TAATTATTCCTCAATTCGCTCAAGATGATACATCTGAGAATCCACTTGT 403
69 roGlyPheLeuAlaProProProPro..... 77
|||||:|||||:
404 ATTCCTTTGGAGGTTCTCCAAAGATAACGTAGAAAGCTGTTCAATTTCA 453
```



```

78 ...AlaProAlaGlnSerProAlaProThrGlnProProLeuProAspAl 93
454 GTGAGACAATCACCAGTAGATGCTTTTGAAGACCTGAAGACCCAGTTT 503
93 aGlyValGlyGluLeuAlaProAspLeuLeuGluGly..... 106
504 GGGAGCAGCAGAGATTCTCTCACCATTATTTATCTGAAGCTGAATCGATG 553
107 .....IleAlaGluAspSerIleAspSer 114
554 GTCTAAAGAGAGAAATGCTGAAGATCTTTGAAGCAAAATGGATAGT 603
115 IleIleValAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnG1 131
604 GTACATGTC.....AAAGATGATTGAATCTCTGGGAGAAATGA 644
131 uGlnPro.....GlnAlaLysValThrArgSerIleValPheValT 145
645 GGTTCCTCTCTTGGCTGGGGCAAAATGTCATGAACATCATAGTAGTTG 694
145 hrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCys 161
695 CTGCAGATGTCTCTCTGTCCTCAAAACAGGTGGCTTGGAGATGTTGCA 744
162 GlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValVa 178
745 GGAGCATTTGCCGAAGGCTTTGGCCAGAGAGACATAGGTCATGGTTGT 794
178 lMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaL 195
795 GGCACCAAGGTATGAAATGCTGTAACCCCAAGATATAGGAGCCGCA 844
195 euTyrThrAlaLysHisLysIleProCysPheGlyGlySerHisGlu 211
845 AATACTACAAGGTTTCAT.....GGCAGGATATGGA 876
212 ValThrPheHisGluTyrArgAspAsnValAspTyrPheValPheValAs 228
877 GTAACATTATTTCCATGCTATATACGAGGTGGGATTTGTTTATGGA 926
228 pHisProSerTyr...HisArgProGlySerLeuTyrGlyAspAsnPheG 244
927 TAGTCCAGACTTCCCTACCAGGGGAATCGTATTATGAGGGAAC..... 972
244 lValaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAla 260
973 .....CGAGTGGATATCTTAAACGCTATGATTTGTTCTGCAAGCAGCT 1017
261 CysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln... 276
1018 GTAGAGGTTCTTGGCATGTTCCATGTGGTGGCTTCTGTTATGGAGATGG 1067
277 .AsnCysMetPheValValAsnAspTyrPheHisAlaSerLeuValProValL 293
1068 TAATTTGGCTTTTCATACGAATGATTGGCATACTGCTCTCTTGGCTGTTT 1117
293 euLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSer 309
1118 ATCTGAAGCAATATATCGTGACAAATGGCTTGTGAATATATGCTCGGTCT 1167
310 ThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerTh 326
1168 GTTCTGGTAAATACACATACAGTACCCAGGCTGCTGGTCTGTAGATGA 1217
326 rTyrProAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGluTyrPv 343
1218 CTTCAAAATTTTGGGCTTGGCGGATCCTACTTGGACCTTTTCAGATTGT 1267
343 alPheProGluTyrAlaArgHisAlaLeuAspLysGlyGlyAlaVal 359
1268 ATGACCCCGTC.....GGAGGTGAACACCTC 1293
360 AsnPheLeuLysGlyAlaValAlaThrAlaAspArgIleValThrValSe 376

```

```

1294 AATATTTTGTCTGGCTGAAGACTGCTGACCGAGTGGTTACTGTTAG 1343
376 rGlnGlyTyrSerTyrGluValThrThrAlaGluGlyGlyGlnGlyLeuA 393
1344 CCATGCTTATGATGGAGCTGAAAACATCAGAAAGTGGTTGGGGCCTAC 1393
393 sngLuLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGly 409
1394 ATGAATATATAAATGAAAGTAACCTGGAAGTTTCAAGGTATTGTTAAATG 1443
410 IleAspIleAsnAspTyrPheProThrThrAspLysCysLeuProHis.. 425
1444 ATTGATGCAAAAGGAGTGGAGCCCGAATTTGATGTGCACCTTAAATCGA 1493
426 .....HistySerValAspLeuSer...GlyLysAlaLysC 438
1494 TGGATACACAAATTTATCTCTAGATCTTAGAGATGGTAAAGCCAGTAT 1543
438 ysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspVal 454
1544 GTAAAGGCTGCTTTCAGCGAGAGTGGTCTGCCTGTTGCTGATAATGA 1593
455 ProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLe 471
1594 CCCATCATTTGATTCATTTGGAAGGTAGACCACCAGAAAGCGTGGATCT 1643
471 uIleLysMetAlaLleProGluLeuMetArgGluAspValGlnPheValM 488
1644 CATTCGCGAGGCCATGCTTGGATTGTTCAGTCATGATGTTCAAGTAGTCA 1693
488 etLeuGlySerGlyAspProIlePheGluGlyTyrMetArgSerThrGlu 504
1694 TGTTAGCAGCGGAGGCAAGACCTTGAGAAATTTACTCAGGAACCTTGAG 1743
505 SerSerTyrLysAspLysPheArgGlyTyrPheValGlyPheSerProVa 521
1744 GGTCAACACAGGGAACAAAGTTAGAGCATGGTTCATTTTCAGTAAAGAT 1793
521 lSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgP 538
1794 GGCATATAGAAATTAACAGCAGTCCGACATCTCATGATGCTTCGAGGT 1843
538 heGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrVal 554
1844 TTGAGCCATCGGATTTGAACAGCTTTACCAATGATGATGGAACCATT 1893
555 ProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAs 571
1894 CCAGTGTGATGCTGTTGGGGCTTTAGAGATACAGTCACTCAATTTGA 1943
571 nProPheGlyAlaLysGlyGluGluGlyThrGlyTyrAlaPheSerProL 588
1944 TCTTTTC.....AACGAGTCTGCTTGGTTGGACCTTCGACAGG 1984
588 euthrValAspLysMetLeuTyrAlaLeuArgThrAlaMetSerThrPhe 604
1985 CAGAGGCAGGAAGCTGATCCATGATTAACCTGCTTGAATACATAC 2034
605 ArgGluHisLysProSerTyrGlyLeuMetLysArgGlyMetThrly 621
2035 TGAATTTACAAAGACAGTTGGAAGGGTCTTCAAAACAAGAGGGATGATGCA 2084
621 sAspHisThrTyrAspHisAlaPro..SerSerThrSerArgSerSer 637
2085 AGATCTTAGCTGGGATAATGCTCTCAGCAATACGAGGATGTCTCTTGTG 2134
638 GlyProSerTyrThrAsnProThrSer 646
2135 CAGCCCAAGTACCAATGGTGTCTTCT 2161

```

seq_name: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-018-418-3

```

alignment_scores:      Length: 784
                        Quality: 1200.50
                        Ratio: 2.686
                        Gaps: 20
Percent Similarity: 57.015 Percent Identity: 37.628

alignment_block:
US-09-6774-824-2 x US-10-018-418-3 ..

Align seg 1/1 to: US-10-018-418-3 from: 1 to: 2842

       7 GlyAlaGlyCysLeuAla.....ProSerValArgLeuArgAl 19
||||| ||||| ||||| |||||
194 GGGCGCGAGCTGTCACTGGCCCGTGGCCGCCGCACGCGCACGGCTCG 243
||||| ||||| ||||| |||||
19 aaSProAlaThrAlaAlaArgAla..... 27
||| ::::: ||||| |||||
244 CGACGGAGGTGTGCCGCGCGCGCGCGGGAAGAAGACGCGAGGGTCG 293
||| ::::: ||||| |||||
28 .....SerAlaCysValValArgAlaAlaArgLeuArgLeuAlaArg 41
||| ::::: ||||| |||||
294 ACNAGCAGCGCGCTCGCGAGGACGCCCCCGGCACGCCGCGTGCGCC 343
||||| ||||| |||||
42 GlyArgTyrValAlaGlu.....LeuSerArgGl 51
::: ||||| |||||
344 GCCACCAAGGTCTCGCGGAGCGGAGGATCCCCTCAAGACGCTCATGCGCA 393
||| ::::: ||||| |||||
51 u.....GlyProAlaAlaArgProAlaGlnGlnGlnLeuA 64
::: ||||| ||||| |||||
394 CGCCCGGAAGTGGCGCGCGCACCGCCGACCAGGACGAGCGCG 443
||| ||||| |||||
64 laProLeuValProGlyPheLeuAlaProPro..... 76
|| ||| ||||| ::::: |||||
444 CCGCTCCA.....CCGAGTATGAACGCGACGCCGCGTGAAACGGTGAAC 487
||| ||||| ||||| |||||
77 .....ProAlaPr 79
||||| ||||| |||||
488 AAATCTACGCGCGCGCGCGACCAAAGACAGCGGGCTGCCCGCAC 537
||||| ||||| ||||| |||||
79 oAlaGlnSerProAlaPro.....ThrClnProProLeuProAspAlaGlv 95
||||| ::::: ||||| ||||| |||||
538 CGCAGCGCGCGCCCATTCGCTGCACCCACAAGAGATTACAGTTGAACGGTG 587
||||| ||||| ||||| |||||
95 alGlyGluLeu.....AlaProAspLeuLeuGluGlyIle 107
::::: ||||| ::::: |||||
588 AAAAACAGCTAACGTCTCGCGCGCGACGAGCATGCCAGGTCGTG 637

```

```

1511 ATACATAACATCGCGACCGAGCGCGTGCAGTAGAATTCCTCGTT 1560
329 pLeuGlyLeuProGluTrpTyrclyAlaLeuGluTrpValPheProG 346
1561 CACCGAGTTCCTGACACTACTCG.....G 1586
346 luTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPheLeu 362
1587 AACACTTCAGACTGTACGACCCCGTGGTGTGAGCACCGCAACTACTTC 1636
363 LysGlyAlaValAlaThrAlaAspArgIleValThrValSerGlnGlyTy 379
1637 GCGCGCGCTGAAGATGCGGACGAGTTCGTGCTGAGCCCGGGTA 1686
379 rSerTrpGluValThrAlaGluGlyGlnGlyLeuAsnGluLeuL 396
1687 CCGTGGGAGCTCAAGACGGTGGAGCGCGCTGGGGCTTCACGACATCA 1736
396 euSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIle 412
1737 TAGCGCAGAACACTGGAAGACCGCGCATCGTCAACGGCATCGACAAC 1786
413 AsnAspTrpAspProThrThrAspLysCysLeuProHis..... 425
1787 ATGGAGTGAACCCCGAGGTGGACCTCCACCTCAAGTCGGACGGCTACAC 1836
426 HistyrSerValAspAspLeu...SerGlyLysAlaLysCysLysAlaG 441
1837 CAACCTTCCTCGGAGCGCTGGACTCGGCAAGCGGAGTGCAGAGGAG 1886
441 luLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIle 457
1887 CCCTCAGCGGAGTGGCGCTGACAGTCCGCGCGCGCTGCCGTGCTC 1936
458 GlyPheIleGlyArgLeuAspTyrcLysGlyIleAspLeuIleLysMe 474
1937 GGCTTCATCGCGCTCGAGCGGCGGAGGCGGTGGAGATCATCGCGGA 1986
474 tAlaIleProGluLeuMetArgLysValGlnPheValMetLeuGlyLys 491
1987 GCGCATGCGCTGGATGCTGTGAGCGGAGGTGCAGCTGCTGTCGGCA 2036
491 erGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyr 507
2037 CCGCGCGCCACGACTGGAGAGCATGTCGGGCACCTTCGAGCGGGAGC 2086
508 LysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisAr 524
2087 CACGACAGGTGCGGGGTGGTGGGTTCCTCCGTGCGCTGGCGCACCG 2136
524 gIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProC 541
2137 GATCAGCGCGCGCGCGAGCGCTCTCATGCCCTCCCGGTTCAGCGCGT 2186
541 ysGlyLeuAsnGlnLeuTyralaMetGlnTyrcGlyThrValProValVal 557
2187 GCGGGTGAACACGCTTACGCCATGCCCTACGCGACCGCTCCCGCTCGT 2236
558 HisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGl 574
2237 CACCGCTGCGGGGTGAGGACACCGCTGCCCGCTTCGACCCCTTC.. 2284
574 yAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThrValA 591
2285 .....AACCACTCCGCGCTCGGGTGGAGTTCGACCGCGCGCGAGCGC 2327
591 spLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHis 607
2328 ACAAGCTGATCGAGCGCTCGGCGACTGCCTCCGCACTACCGGGACTAC 2377
608 LysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisTh 624

```

```

2378 RAGGAGAGCTGGAGGGGCTCCAGGAGCGCGCATGTGCGAGGACTTCAG 2427
624 rTrpAspHisAlaProSerSerThrSerArgSerSerSerGlyProSer 640
2428 CTGGAGGATGCGGCAAGCTCTACGAGGAGCTCTCTCAAGGCCAAGT 2477

seq_name: /cgn2_6/ptodata/2/pna/us60_NEW_COMB.seq:US-60-325-448-1034

seq_documentation_block:
; Sequence 1034, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 1034
; LENGTH: 2412
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-1034

alignment_scores:
Quality: 1198.00 Length: 709
Ratio: 2.680 Gaps: 24
Percent Similarity: 63.047 Percent Identity: 42.313

alignment_block:
US-09-674-824-2 x US-60-325-448-1034 ..
Align seg 1/1 to: US-60-325-448-1034 from: 1 to: 2412

2 AlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeu.. 17
:||||| |||
|||||: ||| :|||
396 AGTCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 445
18 .....ArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysV 31
|||||
446 CAAGTAAGAAGCGGCGCTTCTCAGCGCGCGCGCGCGCGCGCGCTGCC 495
31 alVal.....ArgAlaArgLeuArgLeuAlaArgGlyArg 43
:|||||
496 TCAGGAGACGATCGGTCTACCGCGCGCGCGCGCGCGCGCGCG...CGC 542
44 TyrValAlaGluLeuSerArgGlyProAlaAla.....Ar 56
:|||||
543 CACGCGCGCGTGCACCATACGAAGCTCCCGCGCGCGCGCGCGCGTGA 592
56 gProAlaGlnGlnGlnLeuAlaProPro...LeuValProGlyPhe 71
|||||: ||| |||
593 TCCTTCATCCCTAGACAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 642
72 LeuAlaProProProAlaProAlaGlnSerProAlaProThrGlnPr 88
||||| |||
643 ACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCC 692
88 o...ProLeuProAsp.....AlaGlyV 95
||| |||
693 TCCTCCCAAGCGCTGACAATTCGGAATTCGAGAGGATAAGAGCGCAAG 742
95 alGlyGluLeuAlaPro..... 100
||| |||
743 TTCTTGAGAGTGTCTCGAAGCAAGCGCGCGCGCGCGCGCGCGCGCT 792
101 ..... 104
793 GCGGTAGACAGAGACGTGGGATTTCAAGAAATATTTGATCTGAACGA 842
104 uGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaAlaSerG 121
|||||: |||

```

```
843 ACCGACCCCGGAGGATGCGGATGACGATGATGACTGGGCTGATTCA 892
121 lu...GlnAspSerGluIle...MetAspAlaAsnGluInProGluAla 135
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
893 ATGCTCAGATTCTGAGATCGACACGAGGATGACGATTCGGCCCTTTGGCT 942
136 ..LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTy 151
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
943 GGGGAGATGTCATCAAGCTGATCGTGGTGGCTGCTGAATG..... 984
151 ralaLysSerGlyLeuGlyAspValCysGlySerLeuProIleAlaL 168
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
985 .....GGCTTTGAGATGTTGCAGGTGCTTTACCCAAAGCTT 1021
168 euAlaAlaArgGlyHisArgValMetValValMetProArgTyLeuAsn 184
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1022 TGGCGAGGAGGAGCATCGTGTATGTTGTTGTTGTTACCAAGGTAC..... 1065
185 GlySerSerAspLysAsnTyAlaLysAlaLeuTyThrAlaLysHisI 201
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1066 .....GGTGATTACGCGGAAGCCAGGATGTAGGAATCAGCAA 1103
201 elysIleProCysPheGlyGlySerHisGluValThrPheHisGluT 218
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1104 ATACTAAGGCTGCTGGACAGGATCTGGAAGTGAATATTTCATGCAT 1153
218 vArGAspAsnValAspTrpValPheValAspHisProSerTy...His 233
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1154 TTATCGATGGAGTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1203
234 ArgProGlySerLeuTyGlyAspAsnPheGlyAlaPheGlyAspAsn 250
|||:::|||||:::|||||:::|||||:::|||||:::|||||
1204 CGTCAGGATGACATCTATGGGGGAAC.....AGACAGGAATCAT 1244
250 nPheArgTyThrLeuLeuCystTyAlaAlaCysGluAlaProLeuIle 267
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1245 GAAGCGCATGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1294
267 euGluLeuGlyGlyTyrlIleTyGlyGln...AsnCysMetPheVal 282
|||:::|||||:::|||||:::|||||:::|||||:::|||||
1295 TTCCATGCGGTGGTGGCCCTATGGGATGGCAACTTGGTGTCTCTGCA 1344
283 AsnAspTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTy 299
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1345 AACGATTGGCACACTGCATCTCGCTGTTTATCTGAAGGCATATTACAG 1394
299 gProTyGlyValTyArgAspSerArgSerThrLeuValIleHisAsn 316
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1395 AGCAATGGCATGATCCAGTACACATCGCTCTGCTCTGTGTATACATA 1444
316 euAlaHisGlnGlyValGluProAlaSerThrTyProAspLeuGlyLe 332
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1445 TCGCTTACCAGGGCCGTCGCCAGTAGATGAATCCCTACATGGAATTG 1494
333 ProProGluTrpTyGlyAlaLeuGluTrpValPheProGluTrpAla 349
|||:::|||||:::|||||:::|||||:::|||||:::|||||
1495 CCG..... 1497
349 gArgHisAlaLeuAspLys.....GlyGluA 358
:::|||||:::|||||:::|||||:::|||||:::|||||
1498 .GAGCACTACCTGGATCACTTCAAGCTGTACGACCCCGTCGGCGGAGC 1546
358 laValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValTh 374
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1547 ACCCCAAACATCTTCGGCGGGCCCTCAAGATGGCGACCGGGTGGTACC 1596
375 ValSerGlnGlyTySerTrpGluValThrThrAlaGluGlyGlnG 391
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1597 GTAGCCCCGGCTACCTCTGGGAGCTGACACGACGAGGGCGGCTGGG 1646
391 yLeuAsnGluLeuLeuSerSerValLeuAsnGlyIleValA 408
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1647 CCTCCAGCATCATCAGGGAGAACGACTGGAAGATGAACGGCATCGTA 1696
```

```
408 snGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuPro 424
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1697 ACGCATCGACTACCGGGAGTGAACCCGAGGTGGACGTGCACCTGCAG 1746
425 His.....HisTySerValAspLeu...SerGlyLysAl 436
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1747 TCCGACGGCTACGCCAACTACACCGTGGCTGCTGGACTGCAGCAAGCC 1796
436 aLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgL 453
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1797 CGGTGCAAGCGCGCTGAGCGCGAGCTGGGGCTGGAGTGGCGGACG 1846
453 spValProLeuIleGlyPheIleGlyArgLeuAspTyThrGlnLysGly 469
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1847 ACGTGGCGCTGATCGGCTTCATCGGGCGCTCGACGGGACAGAAAGGTG 1896
470 AspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 486
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1897 GACATCATCGCGCAGCGGATGCGGTGATCGCGGGCAGGACGTGCAGCT 1946
486 eValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSer 503
|||:::|||||:::|||||:::|||||:::|||||:::|||||
1947 GTCTGCTGGCTCCGCGCCCGCCGACCTGGAGGTGATGCTGCAGCGGT 1996
503 hrGluSerSerTyLysAspLysPheArgGlyTrpValGlyPheSerVal 519
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1997 TCGAGGGCGACGACACAGCAAGGTGCGGGTGGGTGGGTTCCTCGGTG 2046
520 ProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 536
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2047 AAGATGGCGCACCGGATCACGCGGGCGCCGACGTGTGTCATGCCGTC 2096
536 rArgPheGluProCysGlyLeuAsnGlnLeuTyThrAlaMetGlnTy 553
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2097 GCGGTTCGAGCGCTCGGCGCTCAACACAGCTCTACGCCATGGCGTACG 2146
553 hrValProValValHisGlyThrGlyLeuArgAspThrValGluThr 569
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2147 CCGTCCCGCTGTCGACCGCTCGGGGCTGAGGGACACCGCTGTTCGGCG 2196
570 PheAsnProPheGlyAlaLysGlyGluGlyThrGlyThrAlaPheSe 586
|||:::|||||:::|||||:::|||||:::|||||:::|||||
2197 TTCGACCGCTTC.....GAGGACACCGCGCTCGGGTGGAGTTCGA 2237
586 rProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSer 603
|||:::|||||:::|||||:::|||||:::|||||:::|||||
2238 CCGCGCCGAGCGCGACAAAGCTCATCGAGCGCTCGGCCCATCGCTCGA 2287
603 hrPheArgGluHisLysProSerTrpGlyLeuMetLysArgGlyMet 619
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2288 CGTACCGCAAGTACAAGGAGAGCTGGAGGGTTTCCAGGTGGCGGCATG 2337
620 ThrLysAspHisThrTrpAspHisAla.ProSerSerThrSerArgSer 636
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2338 TCGCAGGACCTCAGCTGGGACACCGCCGCGAGCTCTACGAGGAGTCT 2387
636 erSerGlyProSerTrpThrAsn 643
|||:::|||||:::|||||:::|||||:::|||||:::|||||
2388 TGTCAAGGCCAAGT...ACCAAT 2407
```

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:US-09-952-677-5

seq_documentation_block:

; Sequence 5, Application US/09952677

; GENERAL INFORMATION:

; APPLICANT: Block, Martina

; Lorz, Horst

; Lutticke, Stephanie

; Walter, Lennart

; Froberg, Claus

; Kossmann, Jens

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES

FROM WHEAT WHICH ARE INVOLVED IN STARCH
SYNTHESIS

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/952,677
FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/196,390
FILING DATE: 19-Nov-1998
APPLICATION NUMBER: DE 196 21 588.9
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: DE 196 36 917.7
FILING DATE: 11-SEP-1996
APPLICATION NUMBER: PCT/EP97/02793
FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:
NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: AGREVO-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 5:

LENGTH: 2825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Triticum aestivum L.
STRAIN: cv. Florida
TISSUE TYPE: ca. 21 d Caryopses
IMMEDIATE SOURCE:
LIBRARY: cDNA library in pBluescript sk (-)
CLONE: pTAS1

FEATURE:
NAME/KEY: CDS
LOCATION: 162..2559
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-952-677-5

alignment_scores:
Quality: 1182.50 Length: 825
Ratio: 2.571 Gaps: 24
Percent Similarity: 55.758 Percent Identity: 36.727

alignment_block:

US-09-674-824-2 x US-09-952-677-5

Align seg 1/1 to: US-09-952-677-5 from: 1 to: 2825

7 GlyAlaGlyCysLeuAla.....ProSerValArgLeuArgal 19
|||||
267 GGGCGCGGAGGTGACCTGGCGCGCGCGCGCGCGCGCGCGCTCG 316
19 aAspProAlaThrAlaAla..... 25
|||

317 CGACGGAGCTGTGGCGCGCTCGCCGCCGGAAGAGGACGCGGGGATCG 366
26 ..ArgAlaSerAlaCysValValArgAlaArgLeuArgLeuAlaArg 41
|||||
367 ACACGCCGCGCGCTCGGTGAGGACGCCCGCGCTCCCGCGGTGGCGCC 416
42 GlyArgTyrValAlaGlu.....LeuSerArgG1 51
|||
417 GCCACCAAGTCCGGAGGACGAGGATCCCGTCAAGACGCTCGACCGCGA 466
51 u.....GlyProAlaAlaArgProAlaGlnGlnGlnL 63
: |||||
467 CGCGCGGAAAGCGCGCGCGCTCCCGCGGAGGAGGAGGAGCGCG 516
63 euAlaProLeuValProGlyPheLeuAlaProProPro..... 76
|||||
517 CCGCTCCGCGGATGATGACGGCATG.....CCGGTGAACGGC 554
77Pr 77
555 GAGACAAATCTACCGCGCGCGCGGCGGCTAAAGACAGCGGGTGC 604
77 oAlaProAlaGlnSerProAlaPro...ThrGlnProProLeuProAspA 93
|||||
605 CACGCCGCGACGCGCGCGCTCCGTCGACCCAGAACAGAGCACCAGTGA 654
93 laglyValGlyGluLeu.....AlaProAspLeuLeuGlu 105
|||
655 ACGTGAAACAAAGCTAAGCTCGCTCGCGCGCGGACGAGCATAGCCGAG 704
106 GlyileAlaGluAspSerileAspSerileLeVal..... 117
|||
705 GCGCGCGCTCGGATTCGCGAGCTACCATTCATCAGCGACAGCGCGCC 754
117 117
755 GGAGTCCGTTGCCAGCTGAGAAGACGCGCGCTCGTCGGCTCAAAAT 804
118AlaAlaSerGluGlnAspSerGluileMetaspAlaasnGlu 131
|||||
805 TCGAGTCTCGGCGCTCTGCTCCGCGGTCTGACACCTGTCAGCGACGCGGAA 854
132 Gln.....ProGlnAlaLy 136
|||
855 CAAGAAGTGAAGAAGGTGGGTCTGTTGTCGAAGAAGCTCCAAAGCCAA 904
136 sval..... 137
905 GGCTCTTTCGCGCTGTCAGCGCGCTGTACAAAGAAGACCTTTGGGATT 954
137 137
955 TCAAGAAATACATTGTTTCGAGGAGCGCGGTGGAGCCCAAGGATGATGC 1004
137 137
1005 CGGGCTGTCGACATGATCGGGCTCTTTGAACACCCACCAGATCACA 1054
138ThrArgSerIleValPheValThrG 146
|||
1055 CTCCGAGACCTTTGGCAGGGGAGAGTGTATGAACGTGGTCTCGTGGCTG 1104
146 lyGluAlaAlaProTyrAlaLysSerGlyLeuGlyAspValCysGly 162
|||||
1105 CTGAGTGTCTCCCTGGTCAAAACAGGTGGTCTGGAGATGTTCGGGT 1154
163 SerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMe 179
|||||
1155 GCTCTGCCCAAGGCTTTGGCAAGAGAGGACATCGTGTATGTTGTTGTT 1204
179 tProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeu 196
|||||
1205 ACCAAGGTAT.....GGGACTATGAAGAAGCCTACGATGTCG 1242

```
196 yrThrAlaLeuHisIleLysIleProCysPheGlyGlySerHisGluVal 212
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1243 GAGTCGCGAAATACTACAAG.....GCTGCTGCACAGGATATGGAAGTG 1286
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 ThrPhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHis 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1287 AATTATTTCATGCTTATATGATGGAGTTGATTTGTTGTCATTGACGC 1336
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
229 sProSerTyr...HisArgProGlySerLeuTyrGlyAspAsnPheGlyA 245
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1337 TCCTCTCTCCGACACCGCTCAGGAGACATTATGGG.....GGCA 1377
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 laPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCys 261
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1378 GCAGACAGGAAATATGAAGCGCATGATTTGTTCTGCAAGCGCGCTGTT 1427
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
262 GluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln...As 277
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1428 GAGGTTCCATGGCACGCTCCATGCGGCGGTGTCCTTATGGGATGGAAA 1477
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
277 nCysMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuL 294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1478 TCTGGTGTATTGCAATGATTGGCACACGGCACTCCTGCTGCTATC 1527
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
294 euAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThr 310
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1528 TGAAGCATATTACAGGACCATGTTTGATGCAGTACACTCGGTCCATT 1577
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
311 LeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTy 327
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1578 ATGGGTGATACATACATGCTCACACGGCGCGTGGCCCTGTAGATGAATT 1627
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
327 rProAspLeuGlyLeuProProGluTrpTyrGlyAlaLeuGluTrpValP 344
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1628 CCCGTTCCAGGTTGCCTGAGCACTACCTG..... 1658
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
344 heProGluTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsn 360
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1659 .....GAACACTTTCAGACTGTACGACCCCGTGGTGTGAACACGCCAAC 1703
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 PheLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGl 377
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1704 TACTTCGGCGCGCGGCTGAAGATGCGGACGAGGTGTGCTGGTGGAGGCC 1753
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
377 nGlyTyrSerTrpGluValThrAlaGluGlyGlnGlyLeuAsnG 394
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1754 CGGGTACCTGTGGAGCTGAAGACGGTGGAGGGCGGTGGGGCTTCACG 1803
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
394 luLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIle 410
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1804 ACATCATACGGCAGAACGACTGGAGAGACCCGCGGCATCGTCAACGGCATC 1853
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
411 AspIleAsnAspTrpAsnProThrThrAspLysCysLeuProHis..... 425
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1854 GACAACATGGAGTGAAGCCCGAGGTGGAGCCGCCACCTCAAGTCGGAGCG 1903
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
426 .....HisTyrSerValAspLeu...SerGlyLysAlaLysCysL 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1904 CTACACCAACTTCTCCCTGAGGACCTGGACTCCGCGCAAGCGGAGTGCA 1953
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
439 ysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValPro 455
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1954 AGGAGGCGCTGCAGGCGGAGCTGGGCGCTGCAGGTCGCGCGCGAGCTGCG 2003
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
456 LeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIl 472
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2004 CTGCTCGGCTTCATCGGCGCTGAGCGGAGAGAGCGGCTGGAGATCAT 2053
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
472 elysMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetL 489
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2054 CGCGGACGCCATGCTGATCGTGCAGCGAGGACGTCGACGTGGTGTATGC 2103
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
489 euGlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSer 505
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2104 TGGGCAACCGGGGCCACAGCACTGGAGAGCATGTGTCAGCACTTCGACGCG 2153
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
506 SerTyrLysAspLysPheArgGlyTrpValGlyPheSerValProValSe 522
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2154 GAGCACACACAAGGTGCGGGGTGGGTGCGGTCTCCGTGCGCTGGC 2203
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
522 rHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheG 539
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2204 GCACCGGATCACGCGGGGGCGGACGCGCTCTCATGCCCTCCCGGTTCG 2253
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
539 luProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValPro 555
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2254 AGCGGTGCGGGCTGAACCACTCTACGCCATGGCCTACGGCACCGTCCCC 2303
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
556 ValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnPr 572
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2304 GTCGTGACCGCGCTCGCGGCTTCAGGACACCGTGCCTGCGCGCTCGACC 2353
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
572 oPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuT 589
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2354 CTTTC.....AACCACTCCGGGCTCGGGTGGACGTTCCGACCGCGCGC 2394
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
589 hrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArg 605
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2395 AGGCGCACAAAGCTGATCGAGGCGCTCGGGCACTGCCCTCCGCACTACCGA 2444
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
606 GluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAs 622
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2445 GACTTCAAGGAGAGCTGGAGGGCCCTCCAGAGCGCGCATGTGCGCAGGA 2494
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
622 pHISThrTrpAspHisAla.ProSerSerThrSerArgSerSerGly 638
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2495 CTTCAGCTGGGAGCACGCGCCCAAGCTCTACGAGGACGCTCTCGTCAAG 2544
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
639 ProSerTrpThrAsn.....ProThrSe 646
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2545 CCAAGTACCAGTGGTGAACTGCTAGCTGCTGCGCACTCCAGCCCGCATGC 2594
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
646 rCysArgArgGlyLeuGly.....ArgSerLysCysGlu...S 658
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2595 GTGCATGACAGGATGGAATGCTATGCGCAGCAGCAAGTGCATGGAG 2644
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
658 erProSerAlaLeuLysThrSer 665
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2645 GCGCGCATCCGCGAAGTACAGT 2667
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/pna/us10_NEW_COMB.seq:us-10-018-418-5
seq_documentation_block:
; Sequence 5, Application US/10018418
; GENERAL INFORMATION:
; APPLICANT: Morell, Matthew
; APPLICANT: Li, Zhongyi
; APPLICANT: Rahman, Sadequr
; APPLICANT: Appels, Rudolph
; TITLE OF INVENTION: Genes Encoding Wheat Starch Synthases and Uses Therefor
; FILE REFERENCE: 127-01
; CURRENT APPLICATION NUMBER: US/10/018,418
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT AU00/00385
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: AU PQ0052/99
; PRIOR FILING DATE: 1999-04-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2107
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
```

37 ArgArgLeuAlaAArgGlyArgTyrValAlaGluLeuSerArgGluGlyPr 53
88 CAAGAACTGGAAGAGGGGCGCGCTGCTGTCGAA.....GAAGCTCC 128
53 oAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLeuValProG 70
129 A.....AAGCCAAAGGCTCTTTCGCCGCTGCAGCCCGG 163
70 ly.....PheLeuAlaProPro 76
164 CTGTACAAGAAGACCTTTGGGATTTCAGAATAACATTGGTTTCGAGG 213
77 ProAlaProAlaGlnSerProAlaProThrGlnProProLeuProAspAl 93
214 CCCGTGGAGGCCAG.....GATGA 233
93 aGlyValGlyGluLeuAlaProAspLeuLeuGluGlyIleAlaGluA 110
234 TGGC.....CGGGCTGTCGACAGT 253
110 spSerIleAspSerIleValAlaAlaSerGluGlnAspSerGluLe 126
254 AT.....CGGGCTCCTTGAACACACACAGAAT 282
127 MetAspAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValph 143
283 CACGACTCCGACCTTTGGCAGGGAGAATGTC...ATGAACGTGTCGT 329
143 eValThrGlyClyAlaAlaProTyrAlaLysSerGlyClyLeuGlyAspV 160
330 CGTGGCTGCTGAGTGTCTCCCTCGTGCACAAACAGGTGCTGCGAGATG 379
160 alCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMet 176
380 TTGGGGTGTCTGCCCAAGGCTTTGGCAAGAGAGAGACATCGTGTATG 429
177 ValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLy 193
430 GTTGTGTACCAAGGTAT.....GGGACATATGAAGAACCTAC 467
193 sAlaLeuTyrThrAlaLysHisIleLysIleProCysaphGlyGlySerH 210
468 GGATGTCGGAGTCCGAAATACTACAG.....CTGCTGGCAGCAGATA 511
210 iSGLuValThrPhePheHisGluTyrArgAspAsnValAspTrpValPhe 226
512 TGAAGAGTAAATATTCCATGCTTATATCGATGAGTGTGATTTGTGTT 561
227 ValAspHisProSerTyr...HisArgProGlySerLeuTyrGlyAspAs 242
562 ATTGACGCTCCTCTCTCCACACCGAGAGAGACATTTATGGG..... 606
242 nPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrA 259
607 ...GGCAGCAGACAGAAATATTAAGACGCATGATTTTGTCTGCAAGG 652
259 laAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGly 275
653 CCCTGTTGAGTTCATGCGACGTTCATCGCGCGGTGTCCTATGGG 702

```

570 PheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSe 586
|||||
1579 TTCACCCCTTC.....AACCACTCCGGCTCGGGTGGACGTTTCCA 1619
586 rProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSert 603
|||
1620 CCGCCGCCGAGCGCACAAAGCTCATCGAGCGCTCGGGCACTGCCTCCGCA 1669
603 rPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
1670 CCTACCGAGACTCAAGAGAGCTGGAGGCCCTCCAGGACGGCGCATG 1719
620 ThrLysAspHisThrTrpAspHisAla_ProSerSerThrSerArgSers 636
|||
1720 TCGCAGGACTTCAGCTGGAGCACGCCCAAGCTCTACGAGGAGTCT 1769
636 exSerGlyProSerTrpThrAn..... 643
1770 CGTCAAGGCCAAGTACCAGTGGTGAACGCTAGCTGCTAGCCGCTCCAGCC 1819
644 ProThrSerCysArgArgGlyLeuGly.....ArgSerLysCy 656
|||
1820 CCGCATGGCTGATCAGAGTGGAACTGCATTCGCACGCGCAGGAAAGTG 1869
656 sglu...SerProSerAlaLeuLysThrSer 665
|||
1870 CCATGGAGCGCGCATCCGCAAGTACAGT 1900

```

seq_name: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-044-543-5

```

seq_documentation_block:
; Sequence 5, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthase Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches
; FILE REFERENCE: 1144D
; CURRENT APPLICATION NUMBER: US/10/044.543
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2348
; TYPE: DNA
; ORGANISM: Curcuma zedoaria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)...(2105)
US-10-044-543-5

```

```

alignment_scores:
  Quality: 1150.00      Length: 517
  Ratio: 3.125          Gaps: 13
  Percent Similarity: 71.180      Percent Identity: 47.389

```

alignment_block:
US-09-674-824-2 x US-10-044-543-5

Align seg 1/1 to: US-10-044-543-5 from: 1 to: 2348

```

140 SerIleValPheValThrGlyGluAlaAlaProTrpAlaLysSerGlyGI 156
|||||
633 AATATCATATTGCTAGCTGCAGAATCGCACCATGCTCTAAACACAGTGG 682
156 yLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyH 173
|||||
683 GCTTGGAGATGTTGTTGGAGCTTTACCTAAAGCATTTGGCCCAAGAGAG 732
173 isArgValMetValValMetProArgTyrLeuAsnGlySerSerAspLys 189
|||||

```

```

733 ATCGTGTCTAGTGTCTCCAGATAT.....GGA 764
190 AsnTyr.....AlaLysAlaLeuTyrThrAlaLysHisIleLysIlePr 204
|||||
765 AACTATCTCAACCTAAGGAATAGGAATCTTAAAGGTACAAGGTTGA 814
204 oCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgAspA 221
|||
815 T.....GGACAGCATCGGATTAATACTATCATACTACATCGATT 858
221 snValAspTrpValPheValAspHisProSerTyrHisArgProGlySer 237
|||||
859 CTGTGTGATTTGCTTCATCGATAGTCTATTTTCCGCCATATTGGAAT 908
238 ..LeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTy 253
|||
909 GATATATATGGTGAAC.....CGAGTGCACATTTTGAAGAGAA 949
253 rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuG 270
|||
950 GGTATTCTTCGAAAGCAGCAGTTGAGTTCCCTTGCATGTCCTCCATGTG 999
270 yGlyTyrIleTyrGlyGln...AsnCysMetPheValValAsnAspTrp 285
|||||
1000 GTGATTTCTGTTATGGAGATGGGAATTTGGTTTTCATTGCCACGATGG 1049
286 HisAlaSerLeuValProValLeuAlaAlaLysTyrArgProTyrGI 302
|||||
1050 CATACCTCCTTACTTCCAGTTTATTGAAGGCATGTTTCGTGATCGTGG 1099
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
|||||
1100 AATATGACATACGCTCGCTCTCTTGGTTATTCAACAACATTCACAT 1149
319 lnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProProGlu 335
|||||
1150 AGGTGCTGGTCCGCTAGATGACTTCTCATATGTGGATTGCCACATGAT 1199
336 TrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAl 352
|||||
1200 CACATTGACTCGTTTAGA..... 1217
352 aleuAspLys.....GlyGluAlaValAsnPheLeuLysGlyAlav 366
|||||
1218 .CTGGATGATCCTGTGGAGGTGAGCATTTTAACATTTTTCACAGCTGTA 1266
366 alValThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGlu 382
|||
1267 TAAGAGCTGCTGACCGTGTGTTACAGTTAGCCATGGCTATGCTGGAG 1316
383 ValThrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerAr 399
|||
1317 TTAACAACTCTGAAGGTGGTGGGATTCATGAGATCATCAACGAGTG 1366
399 glysserValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpA 416
|||||
1367 CCATTGGAAATTCATGATGTTTAAATGGAATCGATACCCATAGTTGA 1416
416 snProThrThrAspLysCysLeuProHis.....HisTyrSer 428
|||
1417 ATCCAAATTTGACGCTCCTTAAATTTCTGATGTTTACACCACTTACC 1466
429 ValAspAspLeuSer...GlyLysAlaLysCysLysAlaGluLeuGlu 444
|||||
1467 CTGGAACCTTGAATGGGAAGGCCAGTGAAGGCTGCTTTGCCAACG 1516
444 sGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG 461
|||||
1517 AGAGTTGGTCTGCTGTTCTGTCGACGAGTCTTCTTCTGCTTCATTG 1566
461 lyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIlePro 477
|||||

```



```

1567 GGAGATTAGACCACTCAAAAAGGTATAGATCTCATAGCGAGGCCATGCAC 1616
478 GluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPr 494
1617 TGGCTGCTCGGTCAAGATCATACAGATAATCATGCTGGGCACCTGGGAGGCC 1666
494 oilePheGluGlyTrpMetArgSerThrGlySerSerThrValAspLysP 511
1667 AGACCTCGAGGATATGCTTCGAAGATTTTCAAGCTGAGCATCGCGTAAGG 1716
511 heArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAla 527
1717 TCAGGGGATGGGTGGGTCTCAGTGAANAATGGCTCATCGGATCACAGCA 1766
528 GlyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAs 544
1767 GGTGCTGATGCCCTACTCATGCCCTCCAGGTTCGAACCTTTGGATTGAA 1816
544 nGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrG 561
1817 CCAACTTCACGCTATGATGTACGGAACAATTCCTGTGTCATGCAGTAG 1866
561 lyGlyLeuArgAspThrValGlnThrPheAsnProPheGlyAlaLysGly 577
1867 GTGGTCTTCGAGATACTGTGCAACAGTTTGATCCGTTT.....AAT 1907
578 GluGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLe 594
1908 GAGACAGGTTTGGGTGACCTTTGACAGGGCAGGCACATAGATGAT 1957
594 uTPAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611
1958 AGTGGCACTCGGCATTTGCTCTAAACACATATCGGAATTACAAGGAGCT 2007
611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627
2008 GGGTGGGATTCAGAACGAGGGATGATGCAGGACCTCAGTTGGGAGAGT 2057
628 AlaProSerSerThrSerArgSerSerGlyProSerTrpThrAsn 643
2058 GCTGCCGAGCAGCATGAAAGTCTTGTGCTGCCAAGT...ACCAAT 2103
seq_name: /cgn2_6/ptodata/2/pna/us09_NEW_COMB_seq2:US-09-865-419A-41741
seq_documentation_block:
; Sequence 41741, Application US/09865419A
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51935)B
; CURRENT APPLICATION NUMBER: US/09/865,419A
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,063
; NUMBER OF SEQ ID NOS: 54020
; SEQ ID NO 41741
; LENGTH: 1936
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700100789_FLI
US-09-865-419A-41741

```

```

alignment_scores:
  Quality: 1068.50      Length: 467
  Ratio: 3.238          Caps: 12
  Percent Similarity: 70.664  Percent Identity: 48.394
alignment_block:
US-09-674-824-2 x US-09-865-419A-41741

```

```

Align seg 1/1 to: US-09-865-419A-41741 from: 1 to: 1936
207 GlyGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValas 223
59 GGACAGGATTCAGAGTGTACTATTATTTCACTCTTACATTGATGAGTTGA 108
223 pTrpValPheValAspHisProSerTyr...HisArgProGlySerLeut 239
109 TTTTGTATTCTAGAACCCCTCCCTCCGCGCACCGGCACANTATATTT 158
239 yrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeu 255
159 ATGGG.....GGAGAAAGATTGGATATTTTGAAGCGCATGATTTTG 199
256 LeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTy 272
200 TTCTGCAAGCGCGTGTGAGGTTCCATGGTATGCTCCATGTGGCGGTAC 249
272 rIleTyrGlyGln...AsnCysMetPheValValAsnAspTrpHisAlas 288
250 TGTCTATGGTGTAGGCAACTTAGTTTATTGCTAATGATTGGCATACCG 299
288 erLeuValProValLeuLeuAlaLysTyrArgProTyrGlyValTyr 304
300 CACTTCGCTGCTCTATCTAAAGCCCTATTACCGGGACAAATGGTTGTG 349
305 ArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyVa 321
350 CAGTATGCTCGCTCTGTGCTGTGATACACACATGCTCATCAGGTCG 399
321 lGluProAlaSerThrTyrProAspLeuGlyLeuProGluTrpTyrG 338
400 TGCCCTGTAGAGCACTTCGTCAATTTTGACTTGCCT..... 436
338 lyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAsp 354
437 .....GAACACTACATCCGAC 451
355 Lys.....GlyGluAlaValAsnPheLeuLy 363
452 CACTTCAAACTGTATGACAACTTGTGGGGATCATCAGCAACGTTTTC 501
363 sGlyAlaValThrAlaAspArgIleValThrValSerGlnGlyTyrS 380
502 TCGGGGGCTGAAGACCGCAGACCGGGTGTGACCGTTAGCAATGGCTACA 551
380 erTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeu 396
552 TGTGGGAGCTGAAGACTTCGGAAGCGGGTGGGGCTCCACGACATCATA 601
397 SerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAs 413
602 AACCAAGACGAGCTGGAAGCTGAGGGGATCGTGAACGGCATCGACATGAG 651
413 nAspTrpAsnProThrThrAspLysCysLeuProHis.....H 426
652 CGAGTGGAAACCCCGCTGTGGAGCTGCACCTCCACTCCGAGACTACACA 701
426 isTyrSerValAspAspLeu...SerGlyLysAlaLysCysLysAlaGlu 441
702 ACTACACGTTTCGAGACGCTGGACACCGGCAAGCGGAGTGAAGCGCCGC 751
442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleG 458
752 CTGCACGGCGAGCTGGGCTTCAGGTCCGACGACGAGCTGCCACTGATCG 801
458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetA 475
802 GTTCATCGGGCGCTGGACCAAGAGGGGCTGGACATCATCCGCCGACG 851
475 laileProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491

```

US-09-674-824-2 x US-09-873-402A-83353
Align seg 1/1 to: US-09-873-402A-83353 from: 1 to: 660
369 AlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrTh 385
3 GCAGATCGAATCGTCACTGCTCAAGTAAAGGGTTATTATGGGAGGTCAACA 52
385 rAlaGluGlyGlnGlnGlyLeuAsnGlnLeuLeuSerSerArgLysSerV 402
53 TGCTGAAGGTGGACAGGGGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTG 102
402 alLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 418
103 TATTAAACGGAATTGTAATGGAATTGACATTAATGATTGGAACCTGCC 152
419 ThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLY 435
153 ACAGACAAATGTATCCCTCTGCTTATTCTGTGATGACCTCTCTCGAAA 202
435 sAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArg 452
203 GGCCAAATGTAAGGTGCATTCAGAAAGGAGCTGGGTTTACCTATAAGGC 252
452 luAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGly 468
253 CTGATGTTCTCTGATGGCTTTATTGGAAGATTGGATTATCAGAAGGC 302
469 IleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGI 485
303 ATTGATCTCATTCACCTTATCATACCCAGATCTCATCGGGGAAGATTGTCA 352
485 nPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgS 502
353 ATTTGTCATGCTTGGATCTGCTGACCCAGAGCTTGAAGATTGGATGAGAT 402
502 erThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSer 518
403 CTACAGAGTCGATCTTCAAGGATAAATTTCTGGATGGGTGGATTAGT 452
519 ValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPr 535
453 GTTCCAGTTTCCCACCGAATAACTGCCGCTGCGATATATTGTTAAATGCC 502
535 oSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG 552
503 ATCCAGATTCCGAACCTTGTGCTCTCAATCAGCTATATGCTATGCGATG 552
552 lyThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGlu 568
553 GCACAGTTCTCTGTTGTCATCACTGGGGCCCTTAGAGATACCGTGGAG 602
569 ThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPh 585
603 AACTTCAACCCCTTTCCGTTGAGAATGGAGAGCGAGGTACAGGGTGGGCATT 652
585 eSerPro 587
653 CGCACCC 659

852 CGATCCACTGGATCGCGGGGACGAGACGCTGCGATGCTGCTGGGACCC 901
492 GlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLY 508
902 GGGCGGGCCGACCTGGAGGACATGCTCGCGGCTTCGAGTCGGAGCACAG 951
508 sAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArgI 525
952 CGACAAGGTGCGCGGTGGGTGGGTTCCTCGGTGCCCTGGCGCACGCCA 1001
525 leThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCys 541
1002 TCACGGCGGGCGCGGACATCTGCTGATGCGCTCGCGGTTCGAGCCGTGC 1051
542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHI 558
1052 GGGTGAACACAGCTCTAGCGCATGCGGTACGGGACCGTGGCGGTGTGCA 1101
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575
1102 CGCGGTGGGGGGCTCGGGACACGGTGGCGCTTCGACCCGTTCC... 1147
575 laLysGlyGluGlyGlyThrGlyTrpAlaPheSerProLeuThrValAsp 591
1148AACGACACCGGGCTCGGTGGAGTTCGACCGCGCGGAGCGAAC 1192
592 LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLY 608
1193 CGGATGATCGACGCGCTCTCGCACTGCGCTACCGAGTACCGGAACTACAA 1242
608 sProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrT 625
1243 CGAGAGCTGGCGCGGCTCGACGGCGCGGATGGCGGACGACCTCAGCT 1292
625 rpAspHisAla.Pro.....SerSerThrSerArgSerSe 636
1293 GGGACACGCGCGCTGCTGTATGAGGACGCTGCTCAAGCGGAAGTAC 1342
636 rSerGlyPro.....SerTrp....ThrAsnProThrSerCysArgArg 649
1343 CAGTGGTGAGCAATAATATGGCGACGCGCGCTGCTGCTCGCAGG 1391
seq_name: /cgn2_6/ptodata/2/pns/US09_NEW_COMB.seq3:US-09-873-402A-83353
seq_documentation_block:
; Sequence 83353, Application US/09873402A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: Varagona, Marguerite J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51934)B
; CURRENT APPLICATION NUMBER: US/09/873,402A
; CURRENT FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,830
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 90966
; SEQ ID NO 83353
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-223-Q6-K6-E12
US-09-873-402A-83353

alignment_scores:
Quality: 1062.00 Length: 219
Ratio: 5.033 Gaps: 0
Percent Similarity: 96.347 Percent Identity: 89.498
alignment_block:

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 14:56:15 ; Search time 60.77 seconds
(without alignments)
947.638 Million cell updates/sec

Title: US-09-674-824-2
Perfect score: 4044
Sequence: 1 MAATGVGAGCLAPSRLRAD.....SDGSLSVRTAEIRNQLVTL 756

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3313	81.9	610	2 T06280	probable ADPglucos
2	2609.5	64.5	626	2 JQ2322	ADPglucose--starch
3	2591	64.1	622	2 T01414	ADPglucose--starch
4	1992	49.3	641	2 T07668	ADPglucose--starch
5	1227.5	30.4	732	2 T01208	ADPglucose--starch
6	1205.5	29.8	698	2 T01209	ADPglucose--starch
7	1204	29.8	752	2 S61505	ADPglucose--starch
8	1154.5	28.5	788	2 T07667	ADPglucose--starch
9	1128.5	27.9	491	2 T06798	ADPglucose--glycog
10	932.5	23.1	608	2 S43341	probable starch sy
11	884.5	21.9	607	1 YUPOY	ADPglucose--starch
12	883	21.8	608	2 T14731	ADPglucose--starch
13	879	21.7	609	2 JQ0703	ADPglucose--starch
14	877	21.7	609	2 S11481	ADPglucose--starch
15	876.5	21.7	603	2 S61504	ADPglucose--starch
16	870	21.5	605	1 S07314	ADPglucose--starch
17	863	21.3	610	2 F86453	granule-bound star
18	861	21.3	608	2 T10906	ADPglucose--starch
19	858.5	21.2	603	1 YUBHY	ADPglucose--starch
20	843.5	20.9	615	1 YUWTV	ADPglucose--starch
21	762.5	18.9	483	2 C86712	ADPglucose--starch
22	758	18.7	322	2 T07924	ADPglucose--starch
23	756	18.7	484	2 S40051	ADPglucose--starch
24	703.5	17.4	486	2 H72321	probable starch sy
25	685.5	17.0	477	1 SVEGCL	glycogen synthase
26	685.5	17.0	477	2 C86009	ADPglucose--starch
27	654.5	16.2	476	2 E83785	glycogen synthase
28	645.5	16.0	484	2 F82165	starch (bacterial
29	626.5	15.5	476	2 C64119	glycogen synthase
					ADPglucose--starch

30	592.5	14.7	477	2 S76496	hypothetical prote
31	590.5	14.6	1071	2 T04926	starch synthase ho
32	589	14.6	463	2 C70363	glycogen synthase
33	579.5	14.3	472	2 B75501	glycogen synthase
34	563	13.9	491	2 S74473	probable ADPglucos
35	556	13.7	1230	2 T07663	soluble starch syn
36	535	13.2	1025	2 H86250	hypothetical prote
37	527.5	13.0	513	2 H83375	probable glycogen
38	509	12.6	441	2 T07926	probable starch sy
39	492	12.2	1674	2 T01265	starch synthase DU
40	419	10.4	474	2 D71469	probable glycogen
41	417	10.3	474	2 A81732	glycogen synthase
42	389.5	9.6	476	2 C72016	glycogen synthase
43	389.5	9.6	476	2 B86609	glycogen synthase
44	346	8.6	521	2 E64500	UDPglucose--starch
45	318.5	7.9	310	2 T01266	starch synthase DU

ALIGNMENTS

RESULT 1

T06280

probable ADPglucose--starch glucosyltransferase (EC 2.4.1.21) precursor - wheat (frag

N:Alternate names: starch synthase

C:Species: Triticum aestivum (common wheat)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999

C:Accession: T06280

R:Block, M.; Loerz, H.; Luetticke, S.

A:Reference number: 215595

A:Accession: T06280

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-610 <BLO>

A:Cross-references: EMBL:U48227; NID:gl373149; PIDN:AAB02197.1; PID:gl373150

A:Experimental source: cv. Florida, endosperm of kernels

C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producin

C:Keywords: glycogen/starch biosynthesis; glucosyltransferase; hexosyltransferase

Query Match 81.9%; Score 3313; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.1e-328;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 EAAPYAKSGGLDVCGLPIALAAARGHRVVMVMPRYLNGSSDKNYAKALYAKHIKIPCF 206

Db 1 EAAPYAKSGGLDVCGLPIALAAARGHRVVMVMPRYLNGSSDKNYAKALYAKHIKIPCF 60

Qy 207 GGSHEVTFEFHYRDNDVAVFDHPSYHRPGSLYGDNGFAGDNGFRYTLACYACEAPLI 266

Db 61 GGSHEVTFEFHYRDNDVAVFDHPSYHRPGSLYGDNGFAGDNGFRYTLACYACEAPLI 120

Qy 267 LELGGYIYGQCMFVYNDVHSLVPLAAKRYPGYVYRDSRSLVTHNLAHQGVPEAST 326

Db 121 LELGGYIYGQCMFVYNDVHSLVPLAAKRYPGYVYRDSRSLVTHNLAHQGVPEAST 180

Qy 327 YPDGLGLPEWYGALEWFPPEWARHDLKGEAVNFKGAVVTADRIVTVSOGYSWEVTTA 386

Db 181 YPDGLGLPEWYGALEWFPPEWARHDLKGEAVNFKGAVVTADRIVTVSOGYSWEVTTA 240

Qy 387 EGGGGLNELLSSRSKSVLNGVINGIDINDWNTTOKCLPHHYSDVDDLKSKAKCAELQKEL 446

Db 241 EGGGGLNELLSSRSKSVLNGVINGIDINDWNTTOKCLPHHYSDVDDLKSKAKCAELQKEL 300

Qy 447 GLPVREDVPLIGFTGRLDYQKIDLIKMAIPELMREDVQFVYMLGSGDPIFEGWMRSTESS 506

Db 301 GLPVREDVPLIGFTGRLDYQKIDLIKMAIPELMREDVQFVYMLGSGDPIFEGWMRSTESS 360

Qy 507 YKDKFRGWGFSVPVSHRITAGCDILLMPSRFPCGLNQLYAMOYGTVPVYVHTGGGLRDT 566

Db 361 YKDKFRGWGFSVPVSHRITAGCDILLMPSRFPCGLNQLYAMOYGTVPVYVHTGGGLRDT 420


```

Db      294  YGVYKDSRSILVHMLAHQGVPEASTYPDGLGPPENWYGALEWFFPEWARRHALDKGEAVN 353
QY      361  FLKGAVVTTADRIVTVSQGYSEWVTTABGGGOGNELLSRRKSVLNGTVNGIDINDWNPTTD 420
Db      354  FLKGAVVTTADRIVTVSQGYSEWVTTABGGGOGNELLSRRKSVLNGTVNGIDINDWNPTD 413
QY      421  KCLPHHYSVDDLGSKAKCKAELQKELGLPVREDVPLIGFTIGRLDYQKIDLIKMAIPELM 480
Db      414  KICPHYSVDDLGSKAKCKGALQKELGLPIRPDVPPLIGFTIGRLDYQKIDLIQLIIPDLM 473
QY      481  REDVOFVLMGSGDPIFFEGWMRSTESSYKDKFCRWGVSFVPVSHRITAGDCIILLMPSRREP 540
Db      474  REDVOFVLMGSGDPELEDNMRSTESIFKDKFCRWGVSFVPVSHRITAGDCIILLMPSRREP 533
QY      541  CGLNQLYAMQYGTVPVYHGTGGLRDVETNPFCAGKEGGTGWAFSPLTVDKM 593
Db      534  CGLNQLYAMQYGTVPVYHATGGLRDVETNPFCGEGGEGTGWAFPLTENN 586

RESULT 4
T07668
ADPglucose--starch glucosyltransferase (EC 2.4.1.21) SSI precursor - potato
C:Species: Solanum tuberosum (potato)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C:Accession: T07668
R:Abel, G.J.W.
submitted to the EMBL Data Library, January 1997
A:Reference number: Z16081
A:Accession: T07668
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-641 <ABE>
A:Cross-references: EMBL:Y10416
A:Experimental source: cv. Desiree; leaf
C:Genetics:
A:Gene: SSI
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose
A:Pathway: starch synthase
C:Superfamily: starch synthase
C:Keywords: glucosyltransferase; hexosyltransferase

```

[illegible]

QY 399 RKSVLNGIVNGIDINDWNPDTDKCLPH---HYSVDDL-SGKAKCAELQKELGLPVRED 453
 Db 506 SDWKFRGIVNGYDTKDWNPQDAYLTSQGYTNINLTKLTQTKRQCKAALQKELGLPVRED 565
 QY 454 VPLIGFGRDLQKQKIDILKMAPELMREDYQFVLMGSDPIFFGWMRSTESSYKDKFRG 513
 Db 566 VPIISFGRDLQKQKIDILKMAPELMREDYQFVLMGSDPIFFGWMRSTESSYKDKFRG 513
 QY 514 WYGFVSPVSHRTAGCDLILLMPSPFPCGLNOLYAMQYGVVPHGGLRDVTETNPFF 573
 Db 626 WYGFVSPVSHRTAGCDLILLMPSPFPCGLNOLYAMQYGVVPHGGLRDVTETNPFF 573
 QY 574 GAKGEGTGWAFSPITVDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWDHA 628
 Db 686 ---DESGVGWTFDRAEANKLMAALWNLCTYKYDKKSWEGIQERMSQDLSWDNA 737
 RESULT 8
 T07667
 UDPglucose--glycogen glucosyltransferase (EC 2.4.1.11) precursor (clone GT11) - potato
 N:Alternate names: glycogen (starch) synthase
 C:Species: Solanum tuberosum (potato)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T07667
 R:Edwards, E.A.; Marshall, J.; Sidebottom, C.; Visser, R.G.F.; Smith, A.M.; Martin, C.
 Plant J. 8, 283-294, 1995
 A:Title: Biochemical and molecular characterisation of a novel starch synthase from potato
 A:Reference number: Z16079; MUID:95400340
 A:Accession: T07667
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-788 <EDW>
 A:Cross-references: EMBL:X87988; NID:g887648; PIDN:CAA61241.1; PID:g887649
 A:Experimental source: cv. Desiree; tuber
 C:Comment: This protein is present both in the stroma of the plastid and tightly bound to
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 28.5%; Score 1154.5; DB 2; Length 788;
 Best Local Similarity 43.7%; Pred. No. 4.3e-74;
 Matches 253; Conservative 83; Mismatches 194; Indels 49; Gaps 12;

QY 79 PAOSPAQTQPL-----PDAGVCELAPDILLLEGIA-----EDSDSIIVAAS 120
 Db 215 PPETPKSSQETLLDVNSRSLVDVPGKKIQSYMPSLRKSSASHVQORNELEGSSAEAN 274
 QY 121 EQDSEINDANEQPAQV---TRSVFTGEAAPYAKSGGLGDCVGLPDLAALARGHVRVW 177
 Db 275 EETEDPVNIDEKPPPLAGTNVMMIILVASECAPWSKGTGLGDVAGALPKALARRGRVWV 334
 QY 178 VMPRYLNGSSDKNYAKALYAKHIKPCFGSGHEVTFEYRDNDVWV-FV-DHPSYHRPG 236
 Db 335 VAPRYDNYEPQDSG---VRKIYKVD--GQVDVDTYFOALLMDCDFVFIHSHMFRHIGN 388
 QY 237 SLYGDNFGAFGDNQFRYTLCLYACAPLILLEGYIYGQ-NCMFVNDWASHVPLVLLA 295
 Db 389 NIYGGN---RVDILKRWLFCKAAIEVPWHPVCGVCYGDGNLVFIANDWHTALLPAYLK 445
 QY 296 AKYRPYGVYRDSRSTLVINLHQGVPEASTYDPLGLPEWYGALEWVPEWARRHALDK 355
 Db 446 AYRONGINMYTRSLVLIHNIHQGRGLEDSYVDLPHPYMDPFKLYDPV-----G 497
 QY 356 GEAVNFKGAVVTADRVTVSOGYSEWVTTAEGGGLNELLSSRSKSVLNGIDINDW 415
 Db 498 GEHNFIFAAGLKTADRVTVSHGYSNELKTSOGGGLHQIINENDWKLOGVINGIDTKEW 557
 QY 416 NPTTDKCLPH-----HYSVDDL-SGKAKCAELQKELGLPVREDYPLIGFGRDLQKQK 469
 Db 558 NPELDVLHPSRDSGYMNYSLDTLTQTKPCKAALQKELGLPVRRDYPVLIIGFGRDLQKQK 617
 QY 470 DLKMAPELMREDYQFVLMGSDPIFFGWMRSTESSYKDKFRGHWGFSVPVSHRITAGC 529
 Db 618 DLIAEAVPMWQDQVQLVNLGTGRRDLQMLQKQFECQHNKIRGWGFSVKTSHRITAGA 677

QY 530 DILLMPSPFPCGLNOLYAMQYGVVPHGGLRDVTETNPFGAKGEGTGWAFSPIT 589
 Db 678 DILLMPSPFPCGLNOLYAMQYGVVPHGGLRDVTQFPDPLMS---QDMGSPDRAE 734
 QY 590 VDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWDHA 628
 Db 735 ASQLPIRINCLLTYREYKKSWEGIQTRCMTQDLSWDNA 773
 RESULT 9
 T06798
 probable starch synthase (EC 2.4.1.1) - wheat (fragment)
 C:Species: Triticum aestivum (common wheat)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Nov-2000
 C:Accession: T06798
 R:Walter, L.; Loerz, H.; Luetticke, S.T.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z15823
 A:Accession: T06798
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-491 <WAL>
 A:Cross-references: EMBL:U66377; NID:gl620659; PIDN:AAB17085.1; PID:gl620660
 A:Experimental source: cv. Florida; endosperm; 21 days after anthesis
 C:Genetics: SSI
 A:Gene: SSI
 C:Superfamily: starch synthase
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 27.9%; Score 1128.5; DB 2; Length 491;
 Best Local Similarity 47.6%; Pred. No. 1.6e-72;
 Matches 236; Conservative 70; Mismatches 163; Indels 27; Gaps 9;

QY 140 SIIVFTGEAAPYAKSGGLGDCVGLPDLAALARGHVRVWVMPRYLNGSSDKNYAKALYAK 199
 Db 1 NVVVAAECSPCKTGGGLGDVAGALPKALAKRGHVRVMPRYLNGSSDKNYAKALYAK 56
 QY 200 HKIKPCFGSGHEVTFEYRDNDVWV-FV-DHPSY-HRPGSLYGDNFGAFGDNQFRYTLCLY 258
 Db 57 YTK---AAGQDMENVYFAYIDGVDFVTDAPLFRHRQEDYIG---GSRQEIIMKRMILFCK 111
 QY 259 AACEAPLILELGGYIYGQ-NCMFVNDWASHVPLVLAARYPYGVYRDSRSTLVINHLA 317
 Db 112 AAVEVPHVPCGGVPGYGDGNLVFIANDWHTALLPYLKAYYRDHGLMQYTRSIMVHINIA 171
 QY 318 HQGVPEASTYDPLGLPEWYGALEWVPEWARRHALDKGEAVNFKGAVVTADRVTVSQ 377
 Db 172 HQGRPVDFEFPTELPEHYL-----EHFRLYDPVSGEHAHNYFAAGLKADQVVVSP 223
 QY 378 GYSWEVTTAEGGGLNELLSSRSKSVLNGIDINDWNPDTDKCLPH-----HYSVDDL- 432
 Db 224 GYLWELKTVGEGWGLHDIIRNDWKTRGIVNGIDNMENPEVDAHLKSDGYTNFSLRTLD 283
 QY 433 SGKAKCAELQKELGLPVREDYPLIGFGRDLQKQKIDILKMAPELMREDYQFVLMGSG 492
 Db 284 SGKRQCKEALQRELGLQVRAVDPVLLGFLGRDLQKQKVEITADAMPWISQDVQLVMLGTG 343
 QY 493 DPIFEGWMRSTESSYKDKFRGHWGFSVPVSHRITAGCDLILLMPSPFPCGLNOLYAMQY 552
 Db 344 RHDLESMLQHFERGHDKVRGWGFSVRLAHRTITAGADALIMPSPFPCGLNOLYAMAYG 403
 QY 553 TVPVVHGTGLDRDVTETNPFGAKGEGTGWAFSPITVDKMLWALRTAMSTPREHKPSWE 612
 Db 404 TVPVVHAGGLRDVFPDFDPE---NHSGLGWTFDRAEHLKLEALGHCLRTYRDFKESWR 460
 QY 613 GLMKRGMTKDHWDHA 628
 Db 461 ALQERMSQDSWEHA 476

RESULT 10

Db 311 GILEADRVLTVPYAAEELISCIARGCELDNIM--RLTGITIVNGMDVSEWDSKDKYI 368
Qy 424 PHHY-SVDDLSGKAKAELOKELGCLPVREDVPLIGFIRGLDYOKGIDLIKMAIPELMRE 482
Db 369 TAKYDATTATIEAKALKALQAGLPVDRKIPLIATFIAGLEEKQKGPVMAAAIPELMQE 428
Qy 483 DVQFVLMGSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFPCG 542
Db 429 DVQIVLLGTGKKFEKLLKSEKYPGKRVAVKFNAPLAHLIMAGADVLAAPSRFPCG 488
Qy 543 LNQLYAMQYGVVPHVHTGGLRDIVETNPFKAGKEEGTGHAFSPLTVVD-----K 592
Db 489 LIQLQGMRYGTFCACASTGGLVDTVI-----EGTGFHMGRLSVDCKVPEPSDKK 539
Qy 593 MLWALRTAMSTFREHKPSWEGMLKRGMTKDHTWDHAPSSTSRSSSGP--SWTN 643
Db 540 VAATLKRAIKVV--GTPAYEEMVRNQMQLSW-----KGPKNWEN 579
RESULT 15
S61504
UDPglucose--starch glucosyltransferase (EC 2.4.1.11) Isoform I precursor - garden pea
N:Alternate names: glycogen(starch) synthase
C:Species: Pisum sativum (garden pea)
C:Date: 20-Jul-1996 #sequence_revision 07-Feb-1997 #text_change 18-Jun-1999
R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
C:Accession: S11481; S22519; S30485; PC2190; JQ2224
Nucleic Acids Res. 18, 5898, 1990
A:Title: Nucleotide sequence of rice waxy gene.
A:Reference number: JQ0703; MUID:91016948
A:Accession: S11481
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-609 <WAN>
A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA3732.1; PID:g577599
R:Okagaki, R.J.
Plant Mol. Biol. 19, 513-516, 1992
A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.
A:Reference number: S22519; MUID:92322986
A:Accession: S22519
A:Molecule type: mRNA
A:Residues: 1-609 <OKA>
A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403
R:Hirano, H.Y.; Sano, Y.
Plant Cell Physiol. 32, 989-997, 1991
A:Title: Molecular characterization of the waxy locus of rice (Oryza sativa).
A:Reference number: S30485
A:Accession: S30485
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-609 <HIR>
A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401
A:Accession: PC2190
A:Molecule type: protein
A:Residues: 78-113 <HIR>
A:Experimental source: Leaf, cDNA pOSLHC2120
C:Comment: This protein is involved in amylose synthesis in the rice endosperm.
C:Genetics:
A:Gene: waxy
A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glucosyltransferase;
F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:78-609/Product: glycogen (starch) synthase #status experimental <MAT>
Query Match 21.7%; Score 877; DB 2; Length 609;
Best Local Similarity 38.5%; Pred. No. 1.7e-54;
Matches 205; Conservative 89; Mismatches 173; Indels 66; Gaps 15;
Qy 140 STVFTVGEAAYAKSGGLGDCGSLPIALAAARGHRVVMVMPRYLNGSSDKYAKALYAK 199
Db 84 NVVYFGAEMAPWSKTGGGLGGLPPAMAARGHRVVMVSPRY-----DQYKDAWDTSV 137
Qy 200 HTKIFCGSGSHVETFEHVRDNVDVDFHPSY-----HRPG-SLYGDNFGA-FGDNQFR 252
Db 138 VAEIKVADRYERVRFFCHYKRGVORFIDHPSFLEKVGKTKGKIYGPDTGVYKDNQMR 197
Qy 253 YTLCLYAAACEAPLILELGGY-----IYQNCMFVYNDVNDVHSLVPLVLLAAKRYPRYGVYRDS 307
Db 198 FSLCQAALEAPRLNPNPKYGTGYEDVVFVNCNDWHTGTPLASYLKNNYQPNIGYRNA 257
Qy 308 RSTLVHNLAHOGVPEASTYPDGLPPENYGALEWVFPWARHARD--KGEAVNELUG 364
Db 258 KVAFCHNISYOGRAFEDYPELNLSEFRSRSDFI-----DGYDTPVEGRKINWMA 310
Qy 365 AVVTADRVTVSQYSWE-VTTAEGGQGLNELLSSKSVLNGIVNGINDNPNPTTKCL 423

Db 311 GILEADRVLTVPYAAEELISCIARGCELDNIM--RLTGITIVNGMDVSEWDSKDKYI 368
Qy 424 PHHY-SVDDLSGKAKAELOKELGCLPVREDVPLIGFIRGLDYOKGIDLIKMAIPELMRE 482
Db 369 TAKYDATTATIEAKALKALQAGLPVDRKIPLIATFIAGLEEKQKGPVMAAAIPELMQE 428
Qy 483 DVQFVLMGSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFPCG 542
Db 429 DVQIVLLGTGKKFEKLLKSEKYPGKRVAVKFNAPLAHLIMAGADVLAAPSRFPCG 488
Qy 543 LNQLYAMQYGVVPHVHTGGLRDIVETNPFKAGKEEGTGHAFSPLTVVD-----K 592
Db 489 LIQLQGMRYGTFCACASTGGLVDTVI-----EGTGFHMGRLSVDCKVPEPSDKK 539
Qy 593 MLWALRTAMSTFREHKPSWEGMLKRGMTKDHTWDHAPSSTSRSSSGP--SWTN 643
Db 540 VAATLKRAIKVV--GTPAYEEMVRNQMQLSW-----KGPKNWEN 579
RESULT 15
S61504
UDPglucose--starch glucosyltransferase (EC 2.4.1.11) Isoform I precursor - garden pea
N:Alternate names: glycogen(starch) synthase
C:Species: Pisum sativum (garden pea)
C:Date: 20-Jul-1996 #sequence_revision 07-Feb-1997 #text_change 18-Jun-1999
R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
C:Accession: S11481; S22519; S30485; PC2190; JQ2224
Nucleic Acids Res. 18, 5898, 1990
A:Title: Nucleotide sequence of rice waxy gene.
A:Reference number: JQ0703; MUID:91016948
A:Accession: S11481
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-609 <WAN>
A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA3732.1; PID:g577599
R:Okagaki, R.J.
Plant Mol. Biol. 19, 513-516, 1992
A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.
A:Reference number: S22519; MUID:92322986
A:Accession: S22519
A:Molecule type: mRNA
A:Residues: 1-609 <OKA>
A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403
R:Hirano, H.Y.; Sano, Y.
Plant Cell Physiol. 32, 989-997, 1991
A:Title: Molecular characterization of the waxy locus of rice (Oryza sativa).
A:Reference number: S30485
A:Accession: S30485
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-609 <HIR>
A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401
A:Accession: PC2190
A:Molecule type: protein
A:Residues: 78-113 <HIR>
A:Experimental source: Leaf, cDNA pOSLHC2120
C:Comment: This protein is involved in amylose synthesis in the rice endosperm.
C:Genetics:
A:Gene: waxy
A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glucosyltransferase;
F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:78-609/Product: glycogen (starch) synthase #status experimental <MAT>
Query Match 21.7%; Score 876.5; DB 2; Length 603;
Best Local Similarity 37.2%; Pred. No. 1.8e-54;
Matches 202; Conservative 102; Mismatches 186; Indels 53; Gaps 15;
Qy 111 SIDSIIVAAEQSDSEIMDANEQPAKVT--RSTVFTVGEAAYAKSGGLGDCGSLPIAL 168
Db 47 SLNKLHVTRATSGSDTSEKSLGKIVCGMSLVFVGAEVGPMWKTGGGLGGLPPVL 106
Qy 169 AARGHRVVMVMPRYLNGSSDKYAKALYAKHIKIPCFGSGSHVETFEHVRDNVDVDFV 228
Db 107 AGNHRVMTVSPRY-----DQYKDAWDTNVLVEKVGDKIETVRFHCFKRGVDRVFD 160
Qy 229 HPSY-----HRPG-SLYGDNFGA-FGDNQFRYTLCLYAAACEAPLILELGG--YI---YQ 276
Db 161 HPLFLERVWKTGSKLXGPKTGIDYRDNLRFSLCQAALEAPRVNLNLSKSYFSGPYGE 220
Qy 277 NCMFVNDVNDVHSLVPLVLLAAKRYPRYGVYRDSRSTLVHNLAHOGVPEASTYPDGLPP 336
Db 221 DVIEVANDWHISALPCYLKSNYKRGYKNAKVAFCFTHNAYQGNRAFSDFSLNLPDEF 280
Qy 337 YGALEWVFPWARHARD-----KGEAVNPLKGVAVTADRVTVSQYSWEVTTTAEFG-G 389
Db 281 RSSDFI-----DGYNKPCGKKIKNMKAGILSDQVFTVSPHYAKELISGEDRG 330
Qy 390 QGLNELLSSKSVLNGIVNGINDNPNPTTKCLPHHYSYVDLS-GRKAKOELQKELGL 448
Db 331 VELDNII--RSTGIIGVNGMDNREWSQTDRIYDVHNETTVEAKPLKGLTQAEIGL 388
Qy 449 PVREDVPLIGFIRGLDYOKGIDLIKMAIPELMREDVQFVLMGSDPIFEGWMRSTESSYK 508

Db 389 PVDSSIPLIGFICKLEQKSDILVEAIAKAFADENQIVVLGTGKKIMEKQIEVLEEKYP 448
QY 509 DKFRGWGFSVPVSHRITACCDILLMPSPREPCGLNQLYAMQYGTVPVHGTGGLRDTVE 568
Db 449 GKAITGTFKNSPLAHKIIAGADFIVIPSPREPCGLVOLHAMPYGTVPVIVSTGGLVDTVK 508
QY 569 T-----FNPFGAKGEEGTGWAFSPLTVDKMLWALRTAMSTFERHKPSPWEGLMKRGMTKD 622
Db 509 EGYTGPHAGFPDVECED-----VDPDDVDKLAATVVRALKTYGTQ--AMKQIILNCMAQN 561
QY 623 HTW 625
Db 562 FSW 564

Search completed: March 28, 2002, 16:20:40
Job time: 5065 sec

THIS PAGE BLANK (USPTO)

Result No.	Query	Score	Match		Length	DB	ID	Description
			Match	%				
1	1	3313	81.9	610	1	UGS2_WHEAT	Q43654	triticum aestivum
2	2	2609.5	64.5	626	1	UGS2_ORYSA	Q40739	oryza sativa
3	3	1992	49.3	641	1	UGS2_SOLTU	P33568	solanum tuberosum
4	4	1204	29.8	752	1	UGS3_PEA	Q43093	pisum sativum
5	5	1154.5	28.5	788	1	UGS3_SOLTU	Q43847	solanum tuberosum
6	6	932.5	23.1	608	1	UGST_MANIS	Q43784	manihot esculenta
7	7	892.5	22.1	608	1	UGST_ANTWA	Q82627	antirrhinum majus
8	8	884.5	21.9	607	1	UGST_SOLTU	Q00775	solanum tuberosum
9	9	883	21.8	608	1	UGST_SORBI	Q43134	sorghum bicolor
10	10	876	21.7	609	1	UGST_ORYSA	Q43092	oryza sativa
11	11	876.5	21.7	603	1	UGST_PEA	P19395	oryza sativa
12	12	872	21.6	609	1	UGST_ORYGL	Q42968	oryza glaberrima
13	13	870	21.5	605	1	UGST_MAIZE	P04713	zea mays
14	14	861	21.3	608	1	UGST_IPOBA	Q42857	ipomoea batatas
15	15	858.5	21.2	603	1	UGST_HORVU	P09842	hordeum vulgare
16	16	843.5	20.9	615	1	UGST_WHEAT	P27736	triticum aestivum
17	17	768.5	19.0	485	1	GLGA_BACST	Q08328	bacillus subtilis
18	18	756	18.7	484	1	GLGA_BACSU	P39125	bacillus subtilis
19	19	696	17.2	480	1	GLGA_AGRVU	P39670	agrobacterium tumefaciens
20	20	685.5	17.0	477	1	GLGA_ECOLI	P08323	escherichia coli
21	21	626.5	15.5	476	1	GLGA_HAETIN	P45179	haemophilus influenzae
22	22	592.5	14.7	477	1	GLGA_SYNY3	P74521	synecococcus
23	23	556	13.7	1230	1	UGS4_SOLTU	Q43846	solanum tuberosum
24	24	346	8.6	521	1	GLGA_METJA	Q59001	methanococcus
25	25	214.5	5.3	2410	1	MOKI_SCHPO	Q9us8	schizosaccharomyces pombe
26	26	193	4.8	1369	1	MOKB_SCHPO	Q9y704	schizosaccharomyces pombe
27	27	177.5	4.4	2397	1	MOKD_SCHPO	Q09854	schizosaccharomyces pombe
28	28	175.5	4.3	2358	1	MOKC_SCHPO	Q9y719	schizosaccharomyces pombe
29	29	173	4.3	2352	1	MOKC_SCHPO	Q9uul4	schizosaccharomyces pombe
30	30	127	3.1	377	1	CTSA_BACSU	P46915	bacillus subtilis
31	31	118.5	2.9	2205	1	POLN_RUBVT	P13889	rubella virus
32	32	118	2.9	351	1	LP5B_RHUME	Q9r9n2	rhizobium meliloti
33	33	115	2.8	875	1	NETR_HUMAN	P56730	homo sapiens

```

Db 61 GGSHEVTFPHRYNDVDFVDPSPHRSGLYDNGFAFGDNQFRYTLCLCYAAEAPLI 120
QY 267 LELGGYIYQNCMFVYNDWHASLVPLVLLAARYPGVYRDSRSTLVHNLAHQGVPEAST 326
Db 121 LELGGYIYQNCMFVYNDWHASLVPLVLLAARYPGVYRDSRSTLVHNLAHQGVPEAST 180
QY 327 YPDGLGPPWYGALEWVFPWARRHALDKGEAVNFKGAVVTADRVTVSQGSYSEVTTA 386
Db 181 YPDGLGPPWYGALEWVFPWARRHALDKGEAVNFKGAVVTADRVTVSQGSYSEVTTA 240
QY 387 EGGQGLNELLSRKSRYLNGVINGIDINDNPTTDKCLPHHYSVDDLSGKAKAELOKEL 446
Db 241 EGGQGLNELLSRKSRYLNGVINGIDINDNPTTDKCLPHHYSVDDLSGKAKAELOKEL 300
QY 447 GLPVREDVPLIGRDLVQKIDILKMAIPELMREDVQFVMLGSDPIFEGWMRSTESS 506
Db 301 GLPVREDVPLIGRDLVQKIDILKMAIPELMREDVQFVMLGSDPIFEGWMRSTESS 360
QY 507 YKDKFRGWGFSVPVSHRITAGCDILMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDT 566
Db 361 YKDKFRGWGFSVPVSHRITAGCDILMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDT 420
QY 567 VETFPNFGAKGEGTGWAFSPLVVDKMLMALRTAMSTFREHKPSWEGLMKRGMTKDTWD 626
Db 421 VETFPNFGAKGEGTGWAFSPLVVDKMLMALRTAMSTFREHKPSWEGLMKRGMTKDTWD 480
QY 627 HAPSSRSRSGSWNTPTSCRGKGRSCSPALKTSSTSSFRGEGYPTLRCPATVE 686
Db 481 HAPSSRSRSGSWNTPTSCRGKGRSCSPALKTSSTSSFRGEGYPTLRCPATVE 540
QY 687 SQACCLWFAGSRITYDCAAAATVAGSGROLQFWGIRKKGCAAGMLTAKHHSDSLSVRVT 746
Db 541 SQACCLWFAGSRITYDCAAAATVAGSGROLQFWGIRKKGCAAGMLTAKHHSDSLSVRVT 600
QY 747 AEIRNLQVLT 756
Db 601 AEIRNLQVLT 610

RESULT 2
UGS2_ORYSA STANDARD: PRT: 626 AA.
AC Q40739;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE SOLUBLE GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11) (SSS).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxId=4530;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.
RC STRAIN=CV. JAPONICA; TISSUE=Seed;
RX MEDLINE=94302151; PubMed=7518089;
RA Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,
RA Kobayashi E., Ohnishi S., Tanaka K.-I., Arai Y.;
RT "Identification, cDNA cloning, and gene expression of soluble starch
RT synthase in rice (Oryza sativa L.) immature seeds.";
RL Plant Physiol. 103:565-573(1993).
CC -|- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] - UDP
CC + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -|- PATHWAY: STARCH BIOSYNTHESIS.
CC -|- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.
CC -|- TISSUE SPECIFICITY: LEAVES AND IMMATURE SEEDS.
CC -|- MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE
CC PURIFIED: RSS1, RSS2 AND RSS3.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D16202; BAA03739.1; -.
CC InterPro: IPR001296; Glycos_transf_1.
CC Pfam: PF00534; Glycos_transf_1;
CC GlycoGen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
CC TRANSIT 1 113 CHLOROPLAST (POTENTIAL).
CC CHAIN 114 626 SOLUBLE GLYCOPEN [STARCH] SYNTHASE, RSS3.
CC CHAIN 122 626 SOLUBLE GLYCOPEN [STARCH] SYNTHASE, RSS1.
CC BINDING 147 147 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 626 AA; 68451 MW; 03E4182507D26658 CRC64;
CC -----
Query Match 64.5%; Score 2609.5; DB 1; Length 626;
Best Local Similarity 82.6%; Pred. No. 2.5e-177; Indels 31; Gaps 9;
Matches 497; Conservative 32; Mismatches 42;
QY 2 AATGVGAGCL-APSVRLRADPATAARASACVVRRLRLARGLRYVAELSRGPAARPAQ 60
Db 5 AGMGGAACLVAPQVR-----PGR-----RLRLQVRR-RCVAELSRDGGSA----H 46
QY 61 QQLAP-----PLVPGFLAP--PPAPAQS--PAPTQPLPDAGVGEIAPDLLLEGIAED 110
Db 47 GLPALAPLQKQVPLTFLVPTSTPAPTQSPAPATPPPLPDGSGVEIEPD--LEGLTED 104
QY 111 SDISIIVAASEDSIMDANEQOAKVTRISIVFTGEAAPYAKSGGLGDCGSLPATAA 170
Db 105 SDIKTIFVASEESEIMDVKEQAQAKVTRSVFVFTGEASPYAKSGGLGDCGSLPATAA 164
QY 171 RGRHVVMVPRYLNGSSDKNYAKALYTAHKIKIPCGGSHEVTFEHEYRDNDVDFVDP 230
Db 165 RGRHVVMVPRYMNALNKNFANAFVTEHKIKIPCGGSHEVTFEHEYRDNDVDFVDP 224
QY 231 SYHRPGSLYGDNFGAFGDNQFRYTLCLCYAAEAPLIELGGYIYQNCMFVYNDWHASLV 290
Db 225 SYHRPGSLYGDNFGAFGDNQFRYTLCLCYAAEAPLIELGGYIYQNCMFVYNDWHASLV 284
QY 291 PVLLAAKYRPGYGVYRDSRSTLVHNLAHQGVPEASTYDGLGPPWYGALEWVFPWARR 350
Db 285 PVLLAAKYRPGYGVYRDSRSTLVHNLAHQGVPEASTYDGLGPPWYGALEWVFPWARR 344
QY 351 HALDKGEAVNFKGAVVTADRVTVSQGSYSEVTTAEGGQGLNELLSRKSRYLNGVINGI 410
Db 345 HALDKGEAVNFKGAVVTADRVTVSQGSYSEVTTAEGGQGLNELLSRKSRYLNGVINGI 404
QY 411 DINDNPTTDKCLPHHYSVDDLSGKAKAELOKELGLPVREDVPLIGRDLVQKID 470
Db 405 DINDNPTTDKCLPHHYSVDDLSGKAKAELOKELGLPVREDVPLIGRDLVQKID 464
QY 471 LKMAIPELMREDVQFVMLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCD 530
Db 465 LKMAIPELMREDVQFVMLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCD 524
QY 531 ILLMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDTVETENPFGAKGEGTGWAFSPLTV 590
Db 525 ILLMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDTVETENPFGAKGEGTGWAFSPLTV 584
QY 591 DK 592
Db 585 EK 586
RESULT 3
UGS2_SOLTU STANDARD: PRT: 641 AA.
ID UGS2_SOLTU
AC P33568;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```


[illegible]

```

Db 285 SLPKALARGHRVWVAPHY-----GNFAEADIDGVRKRYKVA--GQDMVTFYFHTYID 336
QY 221 NVQWVFDHPSYHR-PCSLYGDNFAGNQFRYTLICAAACEAPLILELGGYIYQ-NC 278
Db 337 GVDIVFIDSPFNLESNIYGN--RDLILRMVLFCAAEVPPHVPCCGGICYGDGNL 393
QY 279 MFVVDWHAASLVPVLAAYKRYPVYRDSRSTLVIINLAHQGVPEASTYDGLGLPEWYG 338
Db 394 VFIANDWHTALLPVYKAYYRDHGLMNYTRSLVLIHIAHQGRGPEDEFNTVDLSGNYLD 453
QY 339 ALQWVPEWARRHALDKGAENVLKGAVVTADRIYVTSOGYSWEVTTAEQGGCLNELLS 398
Db 454 LFKWYDPV-----GGEHNFIFAAGLTADRIYVSHGYAWELKTSEGWGLHNLNE 505
QY 399 RKSVLNGIYNGIDINWNTPTDKLPH-----HYSVDDL-SGKAKCAEQLQKELGLPVRED 453
Db 506 SDMKFRGIYNGVDTKWNPOFDAYLTSQGYTNYNLKTLQTKRQCAALQRELGLPVRED 565
QY 454 VPLIGFIRGLDYQKGLDLKMAPELMDREDVQVLMGLSGDPIPEGWMRSTESSYKDKFRG 513
Db 566 VPIISFIRGLDHQKGVLDIAEAPWMSHDVQVLMGLTGRADLEQMLKEFEAQHCKIRS 625
QY 514 WGFSPVSHRITAGDILLMPSRFPCGLNOLYAMOYGVPPVHVGTGGRLDRTVETFNPF 573
Db 626 WVGFSVMAHRIYTAGSDILLMPSRFPCGLNOLYAMYSYGVPPVHVGGVGLDRTVQFENPF 685
QY 574 GAKGEECTGWAFSPLVTDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHDTWDA 628
Db 686 ---DESGVGWTFDRAEANKLMAALWNCLLTYKDYKKSWEQIGRGSQDLSWDNA 737

RESULT 5
ID UGS3_SOLUTU STANDARD; PRT: 788 AA.
AC Q43847;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11) (GBSSI) (GRANULE-
DE BOUND STARCH SYNTHASE II) (FRAGMENT).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 66-75.
RC STRAIN=CV. DESIREE; TISSUE=Tuber;
RX MEDLINE=95400340; PubMed=7670507;
RA Edwards A., Marshall J., Sidebottom C., Visser R.G.F., Smith A.M.,
RA Martin C.;
RT "Biochemical and molecular characterization of a novel starch
RT synthase from potato tubers.";
RL Plant J. 8:283-294(1995).
CC -1- FUNCTION: ACCOUNTS FOR ONLY 10 TO 15% OF THE TOTAL SOLUBLE SPARCH
CC SYNTHASE ACTIVITY IN TUBERS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] -
CC UDP + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -1- PATHWAY: STARCH BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST/AMYLOPLAST, SOLUBLE AND GRANULE-
CC BOUND.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X87988; CAA61241.1; -.

```

```

DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00334; Glycos_transf_1.
KW Glycogen biosynthesis; Transference; Glycosyltransferase;
KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
FT NON_TER 1 1 CHLOROPLAST.
FT TRANSIT <1 65
FT CHAIN 66 788 GLYCOPEN (STARCH) SYNTHASE.
FT BINDING 310 310 UDP-GLUCOSE (BY SIMILARITY).
FT VARIANT 71 71 S -> D.
SQ SEQUENCE 788 AA; 87890 MW; 8D8B90611B862B7B CRC64;

Query Match 28.5%; Score 1154.5; DB 1; Length 788;
Best Local Similarity 43.7%; Pred. No. 4.1e-74;
Matches 253; Conservative 83; Mismatches 194; Indels 49; Gaps 12;

QY 79 PAQSPAPTQPL-----PDAGVGEAPDALLLEGIA-----EDSIDSIIVAAS 120
Db 215 PPTTPASSQETLLDVSRSKSLVDVPGKKIQSYMPSLRKSSASHVQORNELEGSSAEAN 274
QY 121 EQDSEIMDANEQPOAKV---TRSFVETGEAAAPYAKSGGLGVDVCGSLPIALAAARGHRVW 177
Db 275 EETEDPVNIDKPPPLAGTNVNMNIIIVASECAPWSKTGGLGVAGALPKALARRHRVW 334
QY 178 VMPRYLNGSSDKNYAKALYAKHIKIPCFGGSHEVTFHEFYRDNDVDFV-DHPSYHRPG 236
Db 335 VAPRYDNYPEPQDS---VRKIYKVD--GQDVDTYFQALLMDCDFVFIHSMFRIGN 388
QY 237 SLYGDNFGAFGDNQFRYTLICAAACEAPLILELGGYIYQ-NCMFVVDWHAASLVPVLLA 295
Db 389 NIYGN---RVDILKRWLFCAAEVPPHVPCCGGYCGDGNLFTANDWHTALLPAYLK 445
QY 296 AKRYGVYRDSRSTLVIINLAHQGVPEASTYDGLGLPEWYGALWVPEWARRHALDK 355
Db 446 AYYRNGIMNYTRSLVLIHIAHQGRGLEDFSYVDLPPHYMDPPFKLYDPV-----G 497
QY 356 GEAVNELKGAIVTADRIYVTSOGYSWEVTTAEQGGCLNELLSRKSVLNGIYNGIDINW 415
Db 498 GEHNTFAAGLTADRIYVSHGYAWELKTSEGWGLHNLNEQWKLQGLGIVNGIDITKE 557
QY 416 NPTTDRKCLPH-----HYSVDDL-SGKAKCAEQLQKELGLPVREDVPLIGFIRGLDYQKGI 469
Db 558 NPGLDVHLPRSDGYMNYSLDTLQTKRQCAALQKELGLPVREDVPLIGFIRGLDPQKV 617
QY 470 DLIKMAPELMDREDVQVLMGLSGDPIPEGWMRSTESSYKDKFRGWSGVFSPVSHRITAGC 529
Db 618 DLIAEAPWMMGDVQVLMGLTGRDLEQMLRQFECQNDKIRGWGFSYKTSHRITAGA 677
QY 530 DILLMPSRFPCGLNOLYAMOYGVTPVHVGTGGRLDRTVETFNPEFGAKBEGGTGWAFLSPLT 589
Db 678 DILLMPSRFPCALNOLYAMKYGTIPVHVAGGLRDTVQFPDFPLMS---QDWGGPSDRAE 734
QY 590 VDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHDTWDA 628
Db 735 ASQILPRINCLLYREYKKSWEQIGTRCMTQDLSWDNA 773

RESULT 6
ID UGST_MANES STANDARD; PRT: 608 AA.
AC Q43784;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR GBSS.
OS Manihot esculenta (Cassava) (Manioc).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN [1]
RP SEQUENCE FROM N.A.

```

QY	586	SPLTVD	591
	I	:	
Dd	522	GALHVE	527
RESULT	7		
UGST_ANTMA			
ID	UGST_ANTMA	STANDARD;	PRT; 608 AA.
AC	OR82627;		
DT	15-JUL-1999	(Rel. 38, Created)	
DT	15-JUL-1999	(Rel. 38, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).		
GN	WAXY OR GBSS.		
OS	Antirrhinum majus (Garden snapdragon).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
OX	Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.		
OX	NCBI_TaxID=4151;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Leaf;		
RC	Mérida A., Rodriguez J.M., Vincent C., Romero J.M.;		
RA	"The granule-bound starch synthase (GBSS) gene from Antirrhinum majus		
RT	is developmentally and circadian-clock regulated";		
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.		
CC	-1- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] = UDP		
CC	+ (1,4-ALPHA-D-GLUCOSYL)[N+1].		
CC	-1- PATHWAY: STARCH BIOSYNTHESIS.		
CC	-1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.		
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE		
CC	FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/		
CC	or send an email to license@isb-sib.ch).		
DR	EMBL; AJ006293; CAA06958.1; --		
DR	EMBL; AJ006294; CAA06959.1; --		
DR	InterPro; IPR001296; Glycos_transf_1.		
DR	Pfam; PF00534; Glycos_transf_1; 1.		
KW	Glycopen biosynthesis; transferase; Glycosyltransferase;		
FT	Transit peptide; Chloroplast; Starch biosynthesis.		
FT	TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).		
FT	CHAIN 79 608 GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE.		
FT	BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).		
FW	SEQUENCE 608 AA; 66361 MW; 6527D5D3D565B6E0C CRC64;		
SW			
Query Match	22.1%; Score 892.5; DB 1; Length 608;		
Best Local Similarity	39.5%; Pred. No. 1e-55;		
Matches	204; Conservative 96; Mismatches 180; Indels 37; Gaps 15;		
QY	130	NEOQAQVTVTGEAAPYAKSGGLGDVGCSLPALAAHRHYVMVPYRLNG	185
	I	: :	
Dd	69	NGSPLAGKIIIGCTGMNLVFVLAIEVGPWSKTGGLGVDPAMAGNHGRVTMSPY---	125
QY	186	SSDKNYAKALYTAKHKKIPCGGSHEVTFPEHYRDNVDFVDHPSY-----HRPGSLY	239
	I	: :	
Dd	126	--DQYKADWDTSVVWEIKVGDSDIETVRFFHCYKRGVDRFVFDHPIFLEKWGKTKSKIY	162
QY	240	GDNFGA-FGDNQFRYTLILCYAACAEAPLILBL--GGYI---YGONCFMVNDHWASLVPLV	293
	I	: :	
Dd	183	GNAGTYQDNQLRFSLLCOALEAPRVNLNTSSKYFSGPYGEDVVFVANDHTALLPCY	242
QY	294	LAAKYPYGVYRSRSTLVLIHNLAHQGVPEASTYPDGLPPEWYGALEWWFPWARHAL	353
	I	: :	
Dd	243	LKSMYSQKMYLHAFAFCFHNTAYQGFQSSDFCLLNLDQFKSSPDF-FDCGYKEP---	298

```
QY 354 DKGEAVNFKGAVVADRTVTVSQGSWE-VTTAECGQGLNELLSSRKSXVNLGIVNGIDI 412
|| :||:| :||:||||| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 299 VGRKGINNMKAGLESDDRVTVSPYYAMELVSGAERGVELDNVIA--KTSITGIVNGMDT 356
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 413 NDNPNTTKCLPHYSVDD--LSGKAKCKAELOKELGLPVREDVPLIGFTGRIDYQKIDL 471
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 357 QENWPATDKHIDTNYDITVTVMADKPLKALQANAGLPVDKNIPVIGFGRLEEQKGSDI 416
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 472 IKMAIPELMREDOVFYMLGSGDP IFEGWMRSTESSYKDKFRGMVGFSPVSHRITAGCDI 531
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 417 LVAAISKEFGLDVQIIILGTGKKKEEQIQELEVLPYDKARGVAKFNVLPHAMITAGADF 476
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 532 LLMPSEFCGLNQLYAMQYGVVPHGTGGLRDTV-ETFNPF--CAKEEGTGWAFSPL 588
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 477 MLVPSRFEPCGLQLHAMRYGTIPICASTGLVDTVTGTFGHMGAFNVECA--TVDPA 534
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 589 TVDKMLWALRTAMSTFREHKPSMEGLMKRGMTKDHTW 625
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 535 DVOKIATTVERALAAY--GSVAIKEMIQNCMAQDLW 569
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|

RESULT 8
UGST_SOLTU STANDARD; PRT; 607 AA.
AC Q00775; Q43176;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR GBSS.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=79, 7322; PubMed=1886609;
RX MEDLINE=91360072; PubMed=1886609;
RA van der Leij F.R.; Visser R.G.F.; Ponstein A.S.; Jacobsen E.,
RA Feenstra W.J.;
RT "Sequence of the structural gene for granule-bound starch synthase of
RT potato (Solanum tuberosum L.) and evidence for a single point
RT deletion in the anf allele."
RL Mol. Gen. Genet. 228:240-248(1991).
RN [2]
RP REVISIONS.
RA van der Leij F.R.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DONGNONG303;
RA Dai W.L.; Deng W.; Cui M.; Xiu M.; Zhao S.Y.; Wang X.M.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] = UDP
CC + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -I- PATHWAY: STARCH BIOSYNTHESIS.
CC -I- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X58453; CAA41359.1; -.
CC EMBL: X83220; CAA56220.1; -.
CC F01; S16555; YUPOY.
CC InterPro: IPR001296; Glycos_transf_1.
```

```
DR PF00534; Glycos_transf_1; 1.
KW Glycogen biosynthesis; transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 77
FT CHAIN 78 607
FT BINDING 95 95
FT CONFLICT 130 130
FT CONFLICT 338 338
FT CONFLICT 398 398
FT SEQUENCE 607 AA; 66575 MW; 2A377865CF6A650 CRC64;
SQ
Query Match 21.9%; Score 884.5; DB 1; Length 607;
Best Local Similarity 39.1%; Pred. No. 3.9e-55;
Matches 202; Conservative 97; Mismatches 158; Indels 59; Gaps 15;
QY 140 SIVFVTGEAAPKAKSGGLDVGSLPDLAARGHVVMVMPYVNGSSDKYAKALYTAK 199
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 82 NLIFVTEVGPWSKTGGLDGLGGLPPALAAARHVRMTISPRY-----DQYKADWDTSV 135
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 200 HKIPCFGGSHEVTFPEHYRDNDVWVVDHPSY-----HRPGS-LYGDNFGA-FGDNQER 252
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 136 AVEVKVGDSTIEIVRFPHCYKRGVDRVFDHPMFLEKVGWGTGSKYGPAGLDYLDNELR 195
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 253 YTLICVAACEAPLILEL--GGYI---YGONCMFVNDWHASLVPVLLAAKYRPGYVRDS 307
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 196 FSLCQAALEAPKVLNLSNYSFGPYGEDVLFANDWHTALIPCYLKSQYSGRIYLN 255
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 308 RSTLVIHNLAHQGVPEASTYPDLGLPEWYAGALEWYFPEWARRHALD-----KGBAVNF 361
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 256 KVAFCIHNIAYQGRSFSDFFLLNLPDEFGRSDFI-----DGYEKPVKGRKINW 305
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 362 LKGAVTADRIVTYSGYSWEVTTA-EGGQGLNELLSRKSXVNLGIVNGIDINDWPTTD 420
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 306 MKAGIESHRVTVTPSYAQELSAVDKGVGLDSVL--RKTCTIGVINGMDTQEWNPATD 363
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 421 KCLPHYSVDD--LSGKAKCKAELOKELGLPVREDVPLIGFTGRIDYQKIDLTKMAPEL 479
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 364 KYTDVYDITVTMDAKPLKALQAAVGLPVDKKIPILGIFGRLEQKSGDILVAAIHKF 423
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 480 MREDVQFVLMGSDPIPEGWMRSTESSYKDKFRGMVGFSPVSHRITAGCDILLMPSEFE 539
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 424 IGLDVQIVLGTGKKEFEIEQLEVLYPNKAQVAKFNVLPHAMITAGADFLVPSRFE 483
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 540 PCGLNOLYAMQYGVVPHGTGGLRDTVETFNPFAGKEGEGTW-----PLT 589
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 484 PCGLIQHAMRYGTVPICASTGGLVDTVK-----EGYTGFGHGAENVECDVDPAD 534
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 590 VDKMLWALRTAMSTFREHKPSMEGLMKRGMTKDHTW 625
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 535 VLKIVTTVARLAVY--GTLAFAEMIKNCMSELSW 568
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|

RESULT 9
UGST_SORBI STANDARD; PRT; 608 AA.
AC Q43134;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR WX.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. 12311; TISSUE=Seed;
RA Hsing Y.C.; Liu C.; Yu H.; Hsieh J.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] =
CC UDP + (1,4-ALPHA-D-GLUCOSYL)[N+1].
```

01-NOV-1990 (Rel. 16, Created)
01-NOV-1990 (Rel. 16, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
WAXY OR WX.
Oryza sativa (Rice).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzeae; Oryza.
NCBI_TaxID=4530;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Seed; PubMed=1377969;
RC MEDLINE=92322986;
RA Okagaki R.J.;
RT "Nucleotide sequence of a long cDNA from the rice waxy gene.";
RL Plant Mol. Biol. 19:513-516(1992).
[2]
SEQUENCE FROM N.A.
RP STRAIN=CV. JAPONICA TAICHUNG 65; TISSUE=Seedling;
RC Hirano H.Y., Sano Y.;
RA "Molecular characterization of the waxy locus of rice (Oryza
RT sativa).";
RL Plant Cell Physiol. 32:989-997(1991).
[3]
SEQUENCE FROM N.A.
RP STRAIN=CV. JAPONICA, AND CV. HANFENG;
RC MEDLINE=91016948; PubMed=2216792;
RA Wang Z.Y., Wu Z.L., Xing Y.Y., Zheng F.G., Guo X.L., Zhang W.G.,
RT Hong M.M.;
RL "Nucleotide sequence of rice waxy gene.";
RN Nucleic Acids Res. 18:5898-5898(1990).
[4]
SEQUENCE FROM N.A.
RC STRAIN=CV. INDICA;
RA Wang X.Q., Wang Z.Y., Hong M.M.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE OF 153-343 FROM N.A.
RX MEDLINE=91200672; PubMed=2016064;
RA Shimada H., Tada Y.;
RT "Rapid isolation of a rice waxy sequence: a simple PCR method for the
RT analysis of recombinant plasmids from intact Escherichia coli
RT cells.";
RL Gene 98:243-248(1991).
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + {1,4-ALPHA-D-GLUCOSYL}[N] = UDP
CC + {1,4-ALPHA-D-GLUCOSYL}[N+1].
CC -1- PATHWAY: STARCH BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X62134; CAA44065.1; -
DR EMBL; X58228; CAA41186.1; -
DR EMBL; X33694; CAA37732.1; -
DR EMBL; X65183; CAA46294.1; -
DR EMBL; M55039; AAA33918.1; -
DR PIR; JQ0703; JQ0703.
DR PIR; S22519; S22519.
DR PIR; S30485; S30485.
DR PIR; JQ2224; JQ2224.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR GlycoGen biosynthesis; Transferase; Glycosyltransferase;
KW


```

Db      562 PSW 564

!|
RESULT 12
UGST_ORYGL STANDARD; PRT;   609 AA.
AC Q4298;
DT DT    01-NOV-1997 (Rel. 35, Created)
DT DT    01-NOV-1997 (Rel. 35, Last sequence update)
DE DE    15-JUL-1999 (Rel. 38, Last annotation update)
DE GRAMULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY.
OS Oryza glaberrima (African rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
ON NCBI_TaxID=4538;
RX R[1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. GMSI;
RX MEDLINE=92134825; PubMed=1685658;
RA Umeda M., Ohtsubo H., Ohtsubo E.;
RA "Diversification of the rice Wax gene by insertion of mobile DNA
RT elements into introns.";
RL Jpn. J. Genet. 66:569-586(1991).
CC -|- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYOLOSE IN ENDOSPERM.
CC -|- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-BETA-D-GLUCOSYL)[N] = UDP +
CC CC (1,4'-BETA-D-GLUCOSYL)[N+1].
CC -|- PATHWAY: STARCH BIOSYNTHESIS.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensel@isb-sib.ch).
-----
EMBL: D10472; BAA01272.1; -.
InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Glycopen biosynthesis; Transferase; Glycosyltransferase;
FT Transit peptide; Chloroplast; Starch biosynthesis.
FW TRANSIT        1       77     CHLOROPLAST {BY SIMILARITY}.
FT CHAIN          78      609     GRANULE-BOUND GLYCOPEN {STARCH} SYNTHASE.
FT BINDING         97       97     UDP-GLUCOSE {BY SIMILARITY}.
SQ SEQUENCE      609 AA; 66475 MW; C228BBFB9C407FA5 CRC64;
Query Match           21.6%; Score 872; DB 1; Length 609;
Best Local Similarity 38.3%; Pred. No. 3e-54;
Matches 204; Conservative 88; Mismatches 175; Indels 66; Gaps 15;

QY 140 SIPTVTCEAPYAKSGLGVDGSLPTALARGHRVMVMPRYUNGSSDKNYKALTAK 199
   :||| |:||:||||| ||:||| |||||::||| | | | | |
DB 84 NVFVEGAEMPWSKTGGDVGLGPLPAMAANGHVRMVISPY -----DOYKDWDTS 137
   :||| |:||:||||| ||:||| |||||::||| | | | | |
QY 200 HKIPCGGGSHETFFHEHYRDNDVFVDHPSY ----HRPG-SLYGDNGCA-FGDNQFR 252
   :||| |:||:||||| ||:||| |||||::||| | | | | |
DB 138 VAEIKVADRVREVERFFHCYRKGVDREVFDHPSFLFKVGTKGEKIYPGTGVYKDNMR 197
   :||| |:||:||||| ||:||| |||||::||| | | | | |
QY 253 YTLCYAACRAPLIILEGGY -----IYGQNCFMYVNDDHASLVPLLAAKYRTPGYVYRDS 307
   :||| |:||:||||| ||:||| |||||::||| | | | | |
DB 198 FSLLCQAALAPRIELNNPNPFYFGTYGEDVVFCVNDWHGTPLASYLKNNYQPNGIVRNA 257
   :||| |:||:||||| ||:||| |||||::||| | | | | |
QY 308 RSTLVIHNLAHQVEPASTYPDGLGPENWGCALEWFPEWARHALD ---KGEAVNFLKG 364
   :||| |:||:||||| ||:||| |||||::||| | | | | |
DB 258 KVAFCHINIYYGFAPEDYPELNLSRFRRSSFDFI -----DCYDTPVBGRKINMKKA 310

```

```

RESULT 14
UGST_IPOBA
ID      UGST_IPOBA      STANDARD;          PRT;      608 AA.
AC      042857;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      GRANULE-BOUND GLYCOCEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN      WAXY OR SS67.
OS      Ipomoea batatas (Sweet potato) (Batatae).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
OC      Asterales; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OX      NCBI_TaxId=4120;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=CV. TAINONG 57; TISSUE=Tuberous root;
RC      Wang S.J., Yeh K.W., Tsai C.Y.;
RL      Submitted (DEC-1995) to the EMBL/GenBank/DBD databases.
CC      -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC      -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] = UDP
CC      + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC      -1- PATHWAY: STARCH BIOSYNTHESIS.

```

```

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U44126; AAA86423.1; -.
CC DR InterPro: IPR001296; Glycos_transf_1.
CC DR Pfam: PF00534; Glycos_transf_1; 1.
CC KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC KW Transit peptide; Chloroplast; Starch biosynthesis.
CC FT TRANSIT 1 76 CHLOROPLAST (BY SIMILARITY).
CC FT CHAIN 77 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
CC FT BINDING 96 95 UDP-GLUCOSE (BY SIMILARITY).
CC FT SEQUENCE 608 AA; 67000 MW; CCA8FA50A2F69CB0 CRC64;
CC -----
CC 'Query Match 21.3%; Score 861; DB 1; Length 608;
CC Best Local Similarity 38.9%; Pred. No. 1.8e-53;
CC Matches 210; Conservative 77; Mismatches 181; Indels 72; Gaps 17;
CC -----
QY 140 SIVFVTGEAPYAKSGGLGDVCGSLPTALAAGRHVWVMPYRSLNGSSDKNYKALYTAK 199
DB 83 NLVPGVCEEGPWCKTGGDLVGLGPPALAAARGHRVMTCPRY-----DQYKDAWETCV 136
QY 200 HIKPCPGGSHE-VTFHFHYRDNDVWVFDHPVS-----HRPGS-LYGDNFGA-FGDNQF 251
DB 137 VVE-PQVGDRIEPRFPHSYKRGVDRVFDHPHLEKVGKGTGSMGLGPKACKDYKDQL 195
QY 252 RYTLACYAAEAPLILEGG--YI---YGNCMFVNDWHASLPVLLAAKRYPGVYRD 306
DB 196 RFSLLCQAALAPRVLLNLSKYSFGYGEDVFVANDWHHTALLPCYLKTYQSRGIYMN 255
QY 307 SRSTLVTHNLAHOGVEPASTYPDLGLPEWYGALEWFPPEWARRHALD---KGEAVNPLK 363
DB 256 AKVAFCHNIAIYGRFAFDFSLNLDPEYKGSDFDI-----DGVDPKPVKGRKINMK 308
QY 364 GAVYTADRIYTVSGYSWE-VTTAEGQGLNELLSRKSVLNGVINGIDINDWNPPTDKC 422
DB 309 AGIREADRVFTVSPYAKELVSCVSKGVELDNHI--RDCGITGICNGMDTQEWNPATDKY 366
QY 423 LPHHYSVDD--LSGRKAKCAEIQKELGELPVREDVPLIGFGRDLYQKGLIDLKMAPELMR 481
DB 367 LAVKYDITVYVQAARPLKEALQAQVLPVDRNRIPLIGFGRLEQKGSIDILYAASKFIS 426
QY 482 EDVGFVMLGSDPIFEQGMWRSTESSYKDKFRGWGVFSVPVSHRTAGCDILLMPSRPEPC 541
DB 427 MDVQILILGCKKFEQIQIEOLEVWYDPKARGVAKFNVPLAHMITAGADNFLIISREPC 486
QY 542 GLNQLYAMQYGTVPVWVHGTGGLRDTVETFPFGKAGEEGTGWAFSPITVD----- 591
DB 487 GLIQLHAMRYGTPCICASTGGLVDTVK-----EGYTGFGHGAFNVDCEYDVPEDVL 537
QY 592 KMLWALRATSTF-----REHKPSSWEGLMKRGMTKDHDTWHDAPSTSRSSSGP 639
DB 538 KVITTVGRALAIYGTIAETEMIKMKCQSELSWKGPAC-----NWETVLLSLGVAGSEP 590
CC -----
RESULT 15
UGST_HORVU STANDARD; PRT; 603 AA.
ID UGST_HORVU AC P09842;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY.

```


OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=CV. VOGEL-SANGER GOLD; TISSUE=Leaf;
RX MEDLINE=88303345; PubMed=2970062;
RA Rhode W., Becker D., Salamini F.;
RT "Structural analysis of the waxy locus from Hordeum vulgare.";
RN Nucleic Acids Res. 16:7185-7186(1988).
RL [2].
RP SEQUENCE OF 76-89.
RC STRAIN=CV. H354-295-2-5; TISSUE=starchy endosperm;
RX MEDLINE=94170739; PubMed=8125056;
RA Flengstad R.;
RT "Separation of acidic barley endosperm proteins by two-dimensional electrophoresis";
RL Electrophoresis 14:1060-1066(1993).
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -!- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] -> UDP + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -!- PATHWAY: STARCH BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07931; CAA30755.1; -.
DR EMBL: X07932; CAA30756.1; -.
DR PIR: S01727; YUBHY.
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1.1.
KW Glycogen biosynthesis; transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 72 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 73 603 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
FT BINDING 90 90 UDP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 603 AA; 66211 MW; 0B0B3DE6A8217934 CRC64;

Query Match 21.2%; Score 858.5; DB 1; Length 603;
Best Local Similarity 36.0%; Pred. No. 2.7e-53;
Matches 214; Conservative 96; Mismatches 192; Indels 93; Gaps 18;

QY 69 PGF--LAPPPAPA-----QSPAPTQPPIDPACVGEIAPDLLLEGIAEDSIDSIIVA 118
DB 25 PGFGLRPRNPADAALGMRITGASAPKQSKAHRGSRCL-----SVVVS 70

QY 119 ASEQDSEINDANEQPOAKVTRSTVFVTEGAPYAKSGGLDVCGLPIALAAARGHEVMVY 178
DB 71 ATGSG-----MNLVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVMVY 115

QY 179 MPRLNGSSDKNYAKALYTAKHIPKFCGSGHEVTFPEHYRONVDWVFVDHPISY----- 232
DB 116 SPRY-----DQYKDAWDTSVISEIKVADEYERVRFFHCYKRGVDRVFDHPWLEKVRG 169

QY 233 HRPGLYGNFGA-FGDNQFRYTLICYAACEAPLILELGG--YI-----YQNCMFVVNDWH 286
DB 170 KTKETIYGPDACTDYEDNQRFSLCQAALAPRIILNANNPYFSGPYGEDVVFVNCNDWH 229

QY 287 ASLVPVLLAAKVPYGVYRDRSTLVHNLAHQGVPEASTYDPLGLPPPEWYGALEWVPE 346
DB 230 TGLLACYLKSQNSGIYRTAKVAFCIHNIHSYQGRSFDDFAQLNLPDRFKSFDFI--- 286

QY 347 WARRHALDK---GEAVNFKLGAVVATDRIVTVSQGSWEVTTAEG-GQGLNELLSSRKSV 402
DB 287 ----DCYDRPVEGRKINWKKAGILQADKVLTVSPYYAEELISGEARGCCELDNIM--RLTG 340
QY 403 LNSIVNGIDINDWNPPTDKCLPHYSVDD-LSGKAKCAKAELOKELGLPVREDVPLIGFIG 461
DB 341 ITGIVNGMDVSEWDPTKDKFLAVNYDITTALEAKALNKEALQAEVGLPVDRKVPVAVFIG 400
QY 462 RLDYOKGIDLIKMAIPELMR-EDVOFVMLGSGDPIFEGWNRSTESSYKDKFRGWVGSVP 520
DB 401 RLEEQKGPDMVMTAAIPEILKEEDVQILIGTGKKKFEKLLKSMEEKFPCKVRVAVRFNAP 460
QY 521 VSHRITAGCDIILMPSRFPCGLNQLYAMQYGTVPVYVHGTGGRLDRDTVETFPFGAKGEG 580
DB 461 LAHQMMAGADLLAVTSRFPFCGLIQLQGMRYGTPCYCASTGGGLVDIV-----EGK 511
QY 581 TGWAFSPLTVD-----KMLWALRTAMSTFREHKPSWEGLMKRGMTKDHWT 625
DB 512 TGFHMGRLSVDCNVVEPADVKVATTLKRAVKV--GTPAYQBMVKNCMIQDLSW 564

Search completed: March 28, 2002, 16:30:31
Job time: 585 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 16:19:31 ; Search time 100.47 seconds
(without alignments)
1100.646 Million cell updates/sec

Title: US-09-674-824-2
Perfect score: 4044
Sequence: 1 MAATGVGAGCLAPSVRLRAD.....SDGSLSVRTAEIRNLVTL 756

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL17.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp_invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3355	83.0	647	Q9SOH0	Q9sqh0 aegilops ta
2	3351	82.9	647	Q9LEC0	Q9lec0 triticum ae
3	3348	82.8	647	Q9SQG9	Q9sqg9 triticum ae
4	3343	82.7	647	Q9LEB9	Q9leb9 triticum ae
5	3264	80.7	643	Q9MSA3	Q9msa3 hordeum vul
6	2617.5	64.7	623	Q9SSY9	Q9ssy9 oryza sativ
7	2617.5	64.7	626	Q9A206	Q9a206 oryza sativ
8	2591	64.1	622	Q9A064	Q9a064 zea mays (m
9	2562.5	63.4	629	Q9SBL2	Q9sbl2 sorghum bic
10	2038	50.4	652	Q9FNF2	Q9fnf2 arabidopsis
11	2025.5	50.1	575	Q9SEI7	Q9sei7 arabidopsis
12	1258	31.1	792	Q9MAC8	Q9mac8 arabidopsis
13	1245	30.8	630	Q9S365	Q9s365 ipomoea bat
14	1227.5	30.4	732	Q9A899	Q9a899 zea mays (m
15	1205.5	29.8	698	Q9A890	Q9a890 zea mays (m
16	1187	29.4	751	Q9SPH5	Q9sph5 manihot esc
17	1183.5	29.3	576	Q9A926	Q9a926 chlamydomon
18	1178.5	29.1	799	Q9LEE3	Q9lee3 triticum ae
19	1174.5	29.0	799	Q9SPW9	Q9spw9 triticum ae

20	1165	28.8	798	10	Q9LEE2	Q9lee2 triticum ae
21	1148.5	28.4	799	10	Q9M466	Q9m466 triticum ae
22	1128.5	27.9	491	10	Q24398	Q24398 triticum ae
23	930.5	23.0	606	10	Q9XIS6	Q9xis6 phaseolus v
24	890	22.0	708	10	Q64925	Q64925 chlamydomon
25	887.5	21.9	604	10	Q9PRO3	Q9fr03 perilla fru
26	884	21.9	607	10	Q9ZSQ5	Q9zsq5 astragalus
27	868.5	21.5	604	10	Q9SLS6	Q9sls6 triticum tu
28	866.5	21.4	574	10	Q9SYU0	Q9syu0 triticum ae
29	863	21.3	610	10	Q9MAQ0	Q9maq0 arabidopsis
30	861	21.3	605	10	Q9SXX3	Q9sxx3 triticum ae
31	861	21.3	605	10	Q9FUU6	Q9fuu6 triticum ae
32	860	21.3	604	10	Q9SLS9	Q9sls9 triticum tu
33	859	21.2	604	10	Q9SLS7	Q9sls7 triticum tu
34	857	21.2	604	10	Q9S7N5	Q9s7n5 triticum ae
35	854.5	21.1	606	10	Q43012	Q43012 oryza sativ
36	852.5	21.1	605	10	Q9SQ51	Q9sq51 aegilops sp
37	848	21.0	605	10	Q9SQ52	Q9sq52 triticum mo
38	847	20.9	605	10	Q9SLS8	Q9sls8 triticum tu
39	844	20.9	604	10	Q9SXK4	Q9sxx4 triticum ae
40	826	20.4	599	10	Q9SQ58	Q9sq58 triticum ae
41	762.5	18.9	483	2	Q9CHM9	Q9chm9 lactococcus
42	760	18.8	565	10	Q9XEN9	Q9xen9 triticum ae
43	703.5	17.4	486	2	Q9WZZ7	Q9wzz7 thermotoga
44	654.5	16.2	476	2	Q9KDX6	Q9kdx6 bacillus ha
45	649.5	16.1	651	10	Q64927	Q64927 chlamydomon

ALIGNMENTS

RESULT 1
ID Q9SOH0
AC Q9SOH0 PRELIMINARY; PRT: 647 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE STARCH SYNTHASE I.

OS Aegilops tauschii (Patropyrum tauschii).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Aegilops.
OX NCBI_TaxID=37682;

RP SEQUENCE FROM N.A.
RA Li Z., Rahman S., Kosar-Hashemi B., Mouille G., Appels R.,
RA Morell M.K.;
RT "Cloning and characterization of a gene encoding wheat starch synthase
I.";

RL Theor. Appl. Genet. 98:1208-1216(1999).
DR EMBL: AF091802; AAF03557.1; -
DR InterPro; IPR000531; TonB_boxC.

DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
SQ SEQUENCE 647 AA; 71018 MW; 9D4F791F8CE32728_CRC64;

Query Match 83.0%; Score 3355; DB 10; Length 647;
Best Local Similarity 100.0%; Pred. No. 9.2e-241;
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAATGVGAGCLAPSVRLRADPATARASACVVRLRLRLARLGRVVAELSRGPAARPAQ 60
Db 1 MAATGVGAGCLAPSVRLRADPATARASACVVRLRLRLARLGRVVAELSRGPAARPAQ 60
Qy 61 QQLAPPLVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAESIDSIIVAAS 120
Db 61 QQLAPPLVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAESIDSIIVAAS 120
Qy 121 EQDSEIMDANEQPAKYTRSFVFTGEAPYAKSGGLGCVGSLPIALAAARGHRVMVMP 180
Db 121 EQDSEIMDANEQPAKYTRSFVFTGEAPYAKSGGLGCVGSLPIALAAARGHRVMVMP 180

Db	121	EQDSEIMDANEQPOAKVTRSI	VFVTGEAAPYAKSGGLGVDGSLP	IALAARGHRVVMVMP	180
Qy	181	RYLNGSSDKNYAKALYAKHIK	IPCGGSHVTFHEYRDNDVDFV	DHPSVHRPGLSYG	240
Db	181	RYLNGSSDKNYAKALYAKHIK	IPCGGSHVTFHEYRDNDVDFV	DHPSVHRPGLSYG	240
Qy	241	DNFGAGDGNQFRYTL	LLCYAACAPLILELGGY	TYGQNCMFVNDWHSASLVPVLLAAKYRP	300
Db	241	DNFGAGDGNQFRYTL	LLCYAACAPLILELGGY	TYGQNCMFVNDWHSASLVPVLLAAKYRP	300
Qy	301	YGYRDSRSTLV	IHNLAHQGVPEPASTY	PDGLPPEWYGALEWVPEWARRHALDKGEAVN	360
Db	301	YGYRDSRSTLV	IHNLAHQGVPEPASTY	PDGLPPEWYGALEWVPEWARRHALDKGEAVN	360
Qy	361	FLKGAVVTADRI	VTVSOGYSWEVTTAEGG	OGNELLSRKSVLNGVINGIDINDWNP	420
Db	361	FLKGAVVTADRI	VTVSOGYSWEVTTAEGG	OGNELLSRKSVLNGVINGIDINDWNP	420
Qy	421	KCLPHHYSVDDLSG	KAKCAELQKELGPVRED	VPVSHRITAGCDILLMPSRFP	480
Db	421	KCLPHHYSVDDLSG	KAKCAELQKELGPVRED	VPVSHRITAGCDILLMPSRFP	480
Qy	481	REDVQFVLMGSDP	IFEGWNRSTESSYKDK	FRGWGFSVPVSHRITAGCDILLMPSRFP	540
Db	481	REDVQFVLMGSDP	IFEGWNRSTESSYKDK	FRGWGFSVPVSHRITAGCDILLMPSRFP	540
Qy	541	CGLNQLYAMQYGT	VPVVGHTGGLRDTVET	FNPFKAGEEGTGNFASPLTVDKMLWALRTA	600
Db	541	CGLNQLYAMQYGT	VPVVGHTGGLRDTVET	FNPFKAGEEGTGNFASPLTVDKMLWALRTA	600
Qy	601	MSTFREHKPSWEG	LMKRGMTKDHTWDHA	628	
Db	601	MSTFREHKPSWEG	LMKRGMTKDHTWDHA	628	
RESULT 2					
Q9LECO					
ID	Q9LECO	PRELIMINARY;	PRT;	647 AA.	
AC	Q9LECO;				
DT	01-OCT-2000 (TrEMBLrel. 15, Created)				
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)				
DE	STARCH SYNTHASE I-1 PRECURSOR (EC 2.4.1.21).				
GN	WSSI-1.				
OS	Triticum aestivum (Wheat).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;				
OC	Triticeae; Triticum.				
OX	NCBI_TaxID=4565;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. FIELDER; TISSUE=DEVELOPING KERNELS;				
RA	Peng M., Hucl P., Chibbar R.N.;				
RT	"Isolation, characterization and expression analysis of starch				
RT	synthase I from wheat (Triticum aestivum L.).";				
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AJ292521; CAB99209.1; -.				
DR	InterPro; IPR001296; Glycos_transf_1.				
DR	InterPro; IPR000531; TonB_boxC.				
DR	Pfam; PF00534; Glycos_transf_1; 1.				
DR	PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.				
KW	Transit peptide; Transferase; Glycosyltransferase.				
FT	TRANSIT				
FT	CHAIN				
FT	SEQUENCE	647 AA;	71004 MW;	E66D20010B1F9C0A	CRC64;
Query Match 82.9%; Score 3351; DB 10; Length 647;					
Best Local Similarity 99.8%; Pred. No. 1.8e-240;					
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
Qy	1	MAATGVGACGLAPS	VRLRADPATAARASACVVR	ARLRRLARGRYVAELSREGPAARPAQ	60

1	MAATGVGACGLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSREGPAARPAQ	60
61	QOLAPPLVPGFLAPPPAPAPQSPAPTQPLPDAGYGELAPDLLLEGIAEDSIDSIIVAAS	120
61	QOLAPPLVPGFLAPPPAPAPQSPAPTQPLPDAGYGELAPDLLLEGIAEDSIDSIIVAAS	120
121	EQDSEIMDANEQPOAKVTRSIVFVTGEAAPYAKSGGLGVDGSLPIALAAARGHRVVMVMP	180
121	EQDSEIMDANEQPOAKVTRSIVFVTGEAAPYAKSGGLGVDGSLPIALAAARGHRVVMVMP	180
181	RYLNGSSDKNYAKALYAKHIKIPCGGSHVTFTEHEYRDNDVDFVDPHPSVHRPGLSYG	240
181	RYLNGSSDKNYAKALYAKHIKIPCGGSHVTFTEHEYRDNDVDFVDPHPSVHRPGLSYG	240
241	DNFGAGDGNQFRYTLILCYAACAPLILELGGYTYGQNCMFVNDWHSASLVPVLLAAKYRP	300
241	DNFGAGDGNQFRYTLILCYAACAPLILELGGYTYGQNCMFVNDWHSASLVPVLLAAKYRP	300
301	YGYRDSRSTLVITHNLAHQGVPEPASTYPDGLPPEWYGALEWVPEWARRHALDKGEAVN	360
301	YGYRDSRSTLVITHNLAHQGVPEPASTYPDGLPPEWYGALEWVPEWARRHALDKGEAVN	360
361	FLKGAVVTADRIVTVSOGYSWEVTTAEGGQGNELLSRSKSVLNGVINGIDINDWNP	420
361	FLKGAVVTADRIVTVSOGYSWEVTTAEGGQGNELLSRSKSVLNGVINGIDINDWNP	420
421	KCLPHHYSVDDLSGKAKCAELQKELGPLVREDVPLIGFGRDLYQKGDILKMAIPELM	480
421	KCLPHHYSVDDLSGKAKCAELQKELGPLVREDVPLIGFGRDLYQKGDILKMAIPELM	480
481	REDVQFVLMGSDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFP	540
481	REDVQFVLMGSDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFP	540
541	CGLNQLYAMQYGTVPVVGHTGGLRDTVETFNPFKAGEEGTGNFASPLTVDKMLWALRTA	600
541	CGLNQLYAMQYGTVPVVGHTGGLRDTVETFNPFKAGEEGTGNFASPLTVDKMLWALRTA	600
601	MSTFREHKPSWEGMLKRGMTKDHTWDHA	628
601	MSTFREHKPSWEGMLKRGMTKDHTWDHA	628
RESULT	3	
Q9SQG9		
ID	Q9SQG9; PRELIMINARY;	PRT: 647 AA.
AC	Q9SQG9;	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE	STARCH SYNTHASE I.	
OS	Triticum aestivum (Wheat).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;	
OC	Triticeae; Triticum.	
OX	NCBI_taxID=4565;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=ENDOSPERM;	
RA	Li Z., Rahman S., Kosar-Hashemi B., Mouille G., Appels R.,	
RA	Morell M.K.;	
RT	"Cloning and characterization of a gene encoding wheat starch synthase	
RT	I.";	
RL	Theor. Appl. Genet. 98:1208-1216(1999).	
DR	EMBL; AF091803; AAD54861.1; -.	
DR	InterPro; IPR000531; TonB_boxC.	
DR	InterPro; IPR001296; Glycos.transf_1.	
DR	Pfam; PF00534; Glycos.transf_1; 1.	
DR	PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.	
SQ	SEQUENCE	647 AA; 71018 MW; C5D078CCF973BD7D CRC64;
Query Match 82.8%; Score 3348; DB 10; Length 647;		


```
RA Gubler F., Li Z., Fieg S., Jacobsen J.V., Morell M.K.;  
RT "Cloning and characterization of a starch synthase I gene (Accession  
RL No. AF234163) from barley (PCR00-047).";  
RL Plant Physiol. 122:1459-1459(2000).  
DR EMBL; AF234163; AAF37876.1; -.  
DR InterPro; IPR001296; Glycos_transf_1.  
DR InterPro; IPR005334; TonB_boxC.  
DR Pfam; PF00534; Glycos_transf_1; 1.  
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.  
SQ SEQUENCE 643 AA; 70516 MW; C3842A78AF5220EB CRC64;  
  
Query Match 80.7%; Score 3264; DB 10; Length 643;  
Best Local Similarity 97.0%; Pred. No. 5.2e-234;  
Matches 609; Conservative 13; Mismatches 2; Indels 4; Gaps 2;  
  
QY 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLRGRYVAELSRGPAARPAQ 60  
DB 1 MAATGVGAGCLAPSVRLRADP--AARATACVVRARLRRLRGRYVAELSRGPAARPA-- 56  
  
QY 61 QQLAPPLVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAA 120  
DB 57 QQLAPPVVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAEDSIDTIIVAA 116  
  
QY 121 EQDSEIMDANEQOAKVTRSTIVFTVGTGAAPYAKSGGLGDCVGLPITALAARGHRVWVMP 180  
DB 117 EQDSEIMDANEQOAKVTRSTIVFTVGTGAAPYAKSGGLGDCVGLPITALAARGHRVWVMP 176  
  
QY 181 RYLNQSGSDKNYAKALYAKHIKIPCGSGSHEVTFHEYRDNVDMVFDHPHYRPGSLYG 240  
DB 177 RYLNQSGSDKNYAKALYAKHIKIPCGSGSHEVTFHEYRDNVDMVFDHPHYRPGSLYG 236  
  
QY 241 DNFGAFGDNQFRYTLACYACEAPLILELGGYIYGQCMFVNDWHASLVPLVLLAARYP 300  
DB 237 DNFGAFGDNQFRYTLACYACEAPLILELGGYIYGQCMFVNDWHASLVPLVLLAARYP 296  
  
QY 301 YGVYRDRSRLTVIHNLAHQGVPEASTYPDGLPPEWYGALEWVPEWARHLDKGEAVN 360  
DB 297 YGVYRDRSRLTVIHNLAHQGVPEASTYPDGLPPEWYGALEWVPEWARHLDKGEAVN 356  
  
QY 361 FLKAVYADRIYTVSGYSWEVTTAEGGQGLNELLSRKSVLNGIYVNGIDINDNPTTD 420  
DB 357 FLKAVYADRIYTVSGYSWEVTTAEGGQGLNELLSRKSVLNGIYVNGIDINDNPTTD 416  
  
QY 421 KCLPHYSVDDLSGKAKCAELQELGLPVREDVPLIGTGRLDYQKIDILKMAIPELM 480  
DB 417 KCLPHYSVDDLSGKAKCAELQELGLPVREDVPLIGTGRLDYQKIDILKMAIPDLM 476  
  
QY 481 REDVQFVMLGSGDPIFGWNRSTESSYKDKFRGWGVSVPVSHRITAGCDILLMPSRFP 540  
DB 477 REDVQFVMLGSGDPIFGWNRSTESSYKDKFRGWGVSVPVSHRITAGCDILLMPSRFP 536  
  
QY 541 CGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEEGTGAFSPVPLTVDKMLWALRTA 600  
DB 537 CGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEEGTGAFSPVPLTVDKMLWALRTA 596  
  
QY 601 MSTFREHKPSWEGLMKRMGTKDHTWDHA 628  
DB 597 ISTRFHKPSWEGLMKRMGTKDHTWDHA 624  
  
RESULT 6  
Q9SSY9 PRELIMINARY; PRT; 623 AA.  
ID Q9SSY9  
AC Q9SSY9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ESTS A0075322(C11109).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoidae; Oryzeae; Oryza.
```

```
OX NCBI_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Nagamura Y., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC  
RL clone: P0681F10.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB026295; BAA81848.1; -.  
DR InterPro; IPR001296; Glycos_transf_1.  
DR InterPro; IPR002114; PTS_HPr_ser.  
DR Pfam; PF00534; Glycos_transf_1; 1.  
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.  
SQ SEQUENCE 623 AA; 68256 MW; 519F319662568335 CRC64;  
  
Query Match 64.7%; Score 2617.5; DB 10; Length 623;  
Best Local Similarity 82.7%; Pred. No. 5.2e-186;  
Matches 498; Conservative 33; Mismatches 40; Indels 31; Gaps 9;  
  
QY 2 AATGVGAGCL-APSVRLRADPATAARASACVVRARLRRLRGRYVAELSRGPAARPAQ 60  
DB 5 AGMGICAAACLVAPQVR----PGR-----RLRLQVRVR-RCVAELSRDGSQA---Q 46  
  
QY 61 QQLAP-----PLVPGFLAP--PPAPAQS--PAPTQPLPDAGVGLAPDLLLEGIAED 110  
DB 47 RPLAPLAPLVKQPVLPVLTFTPTSTPPAPTQSPAPAPTPPPLPDGSGVEIEPD--LEG 104  
  
QY 111 SIDSIIVAASEQDSEIMDANEQOAKVTRSTIVFTVGTGAAPYAKSGGLGDCVGLPITALA 170  
DB 105 SIKDTITFVASEQDSEIMDVKQAQAKVTRSVVFTVGTGAAPYAKSGGLGDCVGLPITALA 164  
  
QY 171 RGRHVWVMPRYUNGSSDKNYAKALYAKHIKIPCGSGSHEVTFHEYRDNVDMVFDHP 230  
DB 165 RGRHVWVMPRYUNGALNKNFANAFYTERKHIKIPCGSGSHEVTFHEYRDSVDVDFDHP 224  
  
QY 231 SYHRPGSLYDNGFAGDNQFRYTLACYACEAPLILELGGYIYGQCMFVNDWHASLV 290  
DB 225 SYHRPGSLYDNGFAGDNQFRYTLACYACEAPLILELGGYIYGQCMFVNDWHASLV 284  
  
QY 291 PVLLAARYPYGVYRDRSRLTVIHNLAHQGVPEASTYPDGLPPEWYGALEWVPEWAR 350  
DB 285 PVLLAARYPYGVYRDRSRLTVIHNLAHQGVPEASTYPDGLPPEWYGALEWVPEWAR 344  
  
QY 351 HALDKGEAVNFLKAVYADRIYTVSGYSWEVTTAEGGQGLNELLSRKSVLNGIYVNGI 410  
DB 345 HALDKGEAVNFLKAVYADRIYTVSGYSWEVTTAEGGQGLNELLSRKSVLNGIYVNGI 404  
  
QY 411 DINDNPTTDKLPVHYSVDDLSGKAKCAELQELGLPVREDVPLIGTGRLDYQKID 470  
DB 405 DINDNPTTDKLPVHYSVDDLSGKAKCAELQELGLPVREDVPLIGTGRLDYQKID 464  
  
QY 471 LIRKMAIPELMREDVQFVMLGSGDPIFGWNRSTESSYKDKFRGWGVSVPVSHRITAGCD 530  
DB 465 LIRKMAIPELMREDVQFVMLGSGDPIFGWNRSTESSYKDKFRGWGVSVPVSHRITAGCD 524  
  
QY 531 ILLMPSRFEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEEGTGAFSPVPLTV 590  
DB 525 ILLMPSRFEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEEGTGAFSPVPLTV 584  
  
QY 591 DK 592  
DB 585 EK 586  
  
RESULT 7  
Q24206 PRELIMINARY; PRT; 626 AA.  
ID Q24206  
AC Q24206;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE SSS1.
```



```
Db 414 KCIPCHYSVDDLSGKAKGALQKELGLPIRPDPVLIGIFGRLDYQKGLDILQLIIPDL 473
QY 481 REDVQVMLGSGDPIPEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILMPSRREP 540
Db 474 REDVQVMLGSGDPELEDWNRSTESIFDKFRGWGFSVPVSHRITAGCDIILMPSRREP 533
QY 541 CGLNQLYAMQYGTVPVVHGTGGLRDVTETFNPFKAGGEGTGWAFSPLTVDKM 593
Db 534 CGLNQLYAMQYGTVPVVHATGGLRDVTENFPENGEGGEGTGWAFAPLTATENN 586

RESULT 9
Q9SBL2 PRELIMINARY; PRT; 629 AA.
ID Q9SBL2;
AC Q9SBL2;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE SOLUBLE STARCH SYNTHASE.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RAFIR 5765-6-1-11-3;
RA Hsieh J.S., Chen M.R., Hsing Y.I.C.;
RT "Molecular cloning of a Sorghum cDNA encoding the soluble starch
synthase SUBSS-";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF168786; AAD45815.2; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1.1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
SQ SEQUENCE 629 AA; 68548 MW; 24443AF594F35DE5 CRC64;

Query Match 63.4%; Score 2562.5; DB 10; Length 629;
Best Local Similarity 81.1%; Pred. No. 6.5e-182;
Matches 489; Conservative 30; Mismatches 57; Indels 27; Gaps 6;

QY 4 TGVGAGCLAPSRLRADPATAARASACV-----VRALKR---RLAGRYVAELSRREG 52
Db 5 SAVGAACL-----VLARAAAGLGLGGRGDRPRFRQVRRRCVAELSRREG 53

QY 53 PA--ARPAQQOQLAPPLVPGFLAPPAPQAPTOPPLPDAGVCELAPDLLLEGTAED 110
Db 54 PAPTPRPLPALAPPLVPAFLA--PFSEPEGEPASTPPPLPDAGLGLG--LQPEGIAEG 110

QY 111 SIDSIIVAAEODSEIMDANEQOAKVTRISIVFVTEAARYAKSGGLGDCGLSPLTALAA 170
Db 111 SIDETVVVASEQSEIIVGGEQARAKVTSIVFVTEGASPYAKSGGLGDCGLSPLVAAA 170

QY 171 RGRVVMVMPRYLNGSSDKNYAKALYTAHITKIPCFGSGHEVTFEYRDNVDWVFVDHP 230
Db 171 RGRVVMVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFEYRSDVSDWVFVDHP 230

QY 231 SYHRPGSILYGNFCAGFNDGFRYTLICAAACEAPLLELGSIYIGQNCMFVNDWHASLV 290
Db 231 SYHRPGNLIGDKFGAGNDGFRYTLICAAACEAPLVLELGSIYIGQNCMFVNDWHASLV 290

QY 291 PVLLAAKRYPGYVYRDSRSTLVTHNLAHQGVPEASTYVDLGLPPEWYGALEWVPEWARR 350
Db 291 PVLLAAKRYPGYVKDSRSILVTHNLAHQGVPEASTYVDLGLPPEWYGALEWVPEWARR 350

QY 351 HALDKGEAVNFKGAVTADRIYVSGYSWEVTTAGGGGLNELLSRKSVLNGIYNGI 410
Db 351 HALDKGEAVNFKGAVTADRIYVSGYSWEVTTAGGGGLNELLSRKSVLNGIYNGI 410

QY 411 DINDWNPPTKCLPHYSVDDLSGKAKGALQKELGLPIRPDPVLIGIFGRLDYQKGLD 470
```

```
Db 411 DINDWNPATDKCIPCHYSVDDLSGKAKCSALQKELGLPIRPEVPLIGIFGRLDYQKGLD 470
QY 471 LIKMAIPELMREDVQVFMGLSGDPIPEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCD 530
Db 471 LIQLIIPHLMRDQVFMGLSGDPELEDWNRSTESDFDKFRGWGFSVPVSHRITAGCD 530
QY 531 ILLMPSRREPCCGLNQLYAMQYGTVPVVHGTGGLRDVTETFNPFKAGGEGTGWAFSPLTV 590
Db 531 ILLMPSRREPCCGLNQLYAMQYGTVPVVHATGGLRDVTENFPENGEGGEGTGWAFAPLT 590
QY 591 DKM 593
Db 591 ENM 593

RESULT 10
Q9FNF2 PRELIMINARY; PRT; 652 AA.
ID Q9FNF2;
AC Q9FNF2;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE SOLUBLE STARCH SYNTHASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned pl clones.";
RL DNA Res. 4:291-300(1997).
DR EMBL; AB006701; BAB10396.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR002114; PTS_HPR_ser.
DR Pfam; PF00534; Glycos_transf_1.1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
SQ SEQUENCE 652 AA; 72098 MW; 91E5069DCD1B2B5B CRC64;
```

```
Query Match 50.4%; Score 2038; DB 10; Length 652;
Best Local Similarity 62.9%; Pred. No. 6.2e-143;
Matches 379; Conservative 81; Mismatches 93; Indels 50; Gaps 7;

QY 43 RYVAELSRREGPAARPAQQOQLAPPLVPGFLAPP-----PAPAQSPAPTAP 88
Db 64 RFTDAERDGS-----VLGFQLTPPGQQTSTSTGEITHHEKKEAIDQI 111

QY 89 PLPDAGV--GELAPDLLLEGTAEDSIDSIIVAAEODSEIMDANEQOAKVTRISIVFTGE 147
Db 112 VMADFGVPGNRAVE---EGAAEVGIPS-----GKAEVNNLNVFTSE 150

QY 148 AAPYAGSGGLGDCGLSPLTALAAARHVRVMVMPRYLNG--SSDKNYAKALYTAHKIPCF 206
Db 151 AAPYAGSGGLGDCGLSPLTALAAARHVRVMVSPRYLNGTAAADKNYARAKDLGIRVTVNCF 210

QY 207 GGSHEVTFEYRDNVDWVFVDHPHSYHRPGSLYGDNFAGFQDNQFRYTLICAAACEAPLI 266
Db 211 GGSQEVSYFHEYRGDGVDFVDFHKSYPHPCNPGYDGSKAGFQDNQFRYTLICAAACEAPLV 270

QY 267 LELGGYIYQNCMFVNDWHASLVPLVLLAAKRYPGYVYRDSRSTLVTHNLAHQGVPEAST 326
Db 271 LPLGGFTYGEKSLFVNDWHAGLVPLVLLAAKRYPGYVYRDSRSTLVTHNLAHQGVPEAST 330

QY 327 YPDLGLPPEWYGALEWVPEWARRHALDKGEAVNFKLGAVVVTTADRIYVTVSGYSWEVTTA 386
```



```

Db 331 YTNLGLPSEWYGAUCVWPTWARTHALDTGEAVNVLKGAIVTSRIIIVSOGYAWIITV 390
QY 387 EGGCGLNELLSRKSVLNGIYNDINWPTDKCLPHHYSVDLSGKAKCAELQKEL 446
Db 391 EGGYGLQDLLSRKSVNGITNGINVDWNPSTDEHIFPHYSADDSVSEKICKMALQKEL 450
QY 447 GLPVREDVPLIFIGRLDYQKIDILKMAIPELMREDVQFVNLGSGDPFIFEGWWRSTESS 506
Db 451 GLPIRPECPIGIFIGRLDYQKIDILQTAGPDLAVDDIQFVNLGSGDPKYESWWRSMEEET 510
QY 507 YKDFRGWVGFSVPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVYVHGTGGLRDT 566
Db 511 YKDFRGWVGFSVPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVYVHGTGGLRDT 570
QY 567 VETENPFCAKGE-ECTGNAFSPLTVDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWT 625
Db 571 VENFNPAEGGAGTGTGWFTPLSKDSMVSAALRAAATYREYKQSWEGLMRGRMTRNYSW 630
QY 626 DHA 628
Db 631 ENA 633

```

```

RESULT 11
Q9SEI7 PRELIMINARY; PRT; 575 AA.
AC 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE SOLUBLE STARCH SYNTHASE.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsids.
RN NCBI_TaxID=3702;
RP SEQUENCE FROM N.A.
RA Lue W.L., Wang S.M., Yu T.S., Chen J.;
RT "Characterization of Arabidopsis soluble starch synthase gene.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121673; AAF24126.1;
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR002114; PTS_HPT_ser.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR PROSITE; PS00589; PTS_HPT_SER; UNKNOWN_1.
SQ SEQUENCE 575 AA; 6330 MW; 937F9F71EA842030 CRC64;

```

```

Query Match 50.1%; Score 2025.5; DB 10; Length 575;
Best Local Similarity 68.4%; Pred. No. 4.4e-142;
Matches 363; Conservative 81; Mismatches 74; Indels 13; Gaps 4;

QY 109 EDSDSIIV-----AAEODSEIMDANEQPAKVRISIVFVTCGEAAPYAKSGGLGD 159
Db 28 KEALDIQVMADFGPGNRAVEEGAAEV--GIPSGKAEVNNLVFTSEAPYKSGTGLGD 85
QY 160 VCGSLPITALAARGHRVNVMPRYLNG--SSDKYAKALYAKHIKIPCFGSGHEVTFHEY 218
Db 86 VCGSLPITALAARGHRVNVISPRYLNGTAADKNYARAKDLGIRVTVNCFGSGQEVSYFHEH 145
QY 219 RDVNDVDFVDPHSPHRSGLYGNFGAFGDNQPRYTLVLAACEAPLILELGGYIYQNC 278
Db 146 RDGVNDVDFVDPHSPHRSGLYGNFGAFGDNQPRYTLVLAACEAPLILELGGYIYQNC 278
QY 279 MFVNDVNDVHSLVPLLAARYKRYGVYRDSRSLVLIHNLAHQGVPEASTYDPLGLPEWYG 338
Db 206 LFLVNDVNDVHSLVPLLAARYKRYGVYRDSRSLVLIHNLAHQGVPEASTYDPLGLPEWYG 338
QY 339 ALEWVFPENARRHALDKGEAVNFKGAVVTADRTIVTSQGSWEVTTAEGGQGLNELLS 398

```

```

Db 266 AVGVWFTWARTHALDTGEAVNVLKGAIVTSRIIIVSOGYAWIITV 325
QY 399 RKSVLNGIYNDINWPTDKCLPHHYSVDLSGKAKCAELQKELGLPVREDVPLIG 458
Db 326 RKSVINIITNGINVDWNPSTDEHIFPHYSADDSVSEKICKMALQKELGPIRPECPIG 385
QY 459 FIGRLDYQKIDILKMAIPELMREDVQFVNLGSGDPFIFEGWWRSTESSYKDKFRGWVGFS 518
Db 386 FIGRLDYQKIDILQTAGPDLAVDDIQFVNLGSGDPKYESWWRSMEEETYRDKFRGWVGFS 445
QY 519 VPSVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVYVHGTGGLRDTVTFNPFKAKE 578
Db 446 VPSVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVYVHGTGGLRDTVTFNPFKAKE 578
QY 579 -EGTGWAFSPLTVDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWT 628
Db 506 GAGTGWFTPLSKDSMVSAALRAAATYREYKQSWEGLMRGRMTRNYSWENA 556

```

```

RESULT 12
Q9MAC8 PRELIMINARY; PRT; 792 AA.
AC 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PUTATIVE GLYCOGEN SYNTHASE.
GN T4P13.13.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsids.
RN NCBI_TaxID=3702;
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T4P13 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC008261; AAF26156.1;
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002114; PTS_HPT_ser.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00589; PTS_HPT_SER; UNKNOWN_1.
SQ SEQUENCE 792 AA; 87592 MW; F8C852E4119EF670 CRC64;

```

```

Query Match 31.1%; Score 1258; DB 10; Length 792;
Best Local Similarity 44.8%; Pred. No. 7.2e-85;
Matches 278; Conservative 79; Mismatches 202; Indels 62; Gaps 16;

QY 46 AELSRREGPAARPAQQOOLA-----PPLVPGFLAPP--PPAPAQSPAPTPPLPDAG--- 94
Db 181 ASVISSTPVSQKPSQSDVATNGKPSVAVSSVDDPPYKPSVMTSPEKTSPTVTFSGFPS 240
QY 95 ---VGLAPDLL-----LEGIAEDSDSIIVAAEODSEIMDANEQ--- 134
Db 241 KSRAGAFWSDDLPSYLTAKAPQTSMTKTEYVEKTPD---VASSETNEPGKDEKPPPLAG 297
QY 135 AKVTRISIVFTVTCGEAAPYAKSGGLGDCVCSLPITALAARGHRVNVMPRYLNGSSDKYAKA 194
Db 298 ANV--NNVILVAEACAPFSKGTGLGDVAGALPKSLARRGHRVNVMPRYLNGSSDKYAKA 352
QY 195 LYTAKHIKIPCGSGHEVTFHEYRDNDVDFVDPHSPY--HREPSLYGDNFGAFGDNQPRY 253
Db 353 LGVRRYKVA--GQDMVEMYHAFIDGVDFVDFIDSPFRHLSNNIYGN--RDLILARM 407
QY 254 TLLCYAAEAPLILELGGYIYQNC--NCEMVNDVNDVHSLVPLLAARYKRYGVYRDSRSLV 312
Db 408 VLFCKAAVEVPWYPCGGVCGYGDGNLAFIANDWHTALLPVYLKAYIRDHGINKYTRSLV 467

```

[illegible]

```

RESULT 13
O65365 PRELIMINARY; PRT; 630 AA.
AC O65365;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE STARCH SYNTHASE.
GN SSSI.
OS Ipomoea batatas (Sweet potato) (Batate).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
ON NCBI_TaxID=4120;
RX [1]
RN SEQUENCE FROM N.A.
RP Harn C.H., Bae J.M., Liu J.R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RR EMBL; AF068834; AAC19119.1; -
DR Mendel; 29714; Ipoba; Sssl; 29714.
DR InterPro; IPR001296; Glycos.transf_1.
DR InterPro; IPR002114; PTS_HPI_ser.
DR Pfam; PF00534; Glycos.transf_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
DR SO SEQUENCE 630 AA: 69961 MW: C55669D761B1A12A CRG64;

```

Query Match	30.8%	Score 1245;	DB 10;	Length 630;
Best Local Similarity	44.8%;	pred. No. 4.7e-84;		
Matches 275;	Conservative 82;	Mismatches 197;	Indels 60;	Gaps 14;
Qy	52	GPARPAAQQQ--LAPPLVPGLA---PPPAPAPQSPAPTO-PP-L---PDAGVGELAPDL	102	
Db	25	GSTERPAVKSKVLSLPSQVTSSTVNSQEPAKATYETVKSSQVMPDLVDVSQKNVTSFSREI	84	
Qy	103	LEGTAEISDIIIVAASQDSQSEIMANEQPAKV-----	141	
Db	85	LL---SEYPPSSSMASASTLEDEKRDHRESSKETDVGTEDPVNBDLKPPPLACTNVNV	141	
Qy	142	VFVTGEAAPAKSGGLGDCVSLTALAARGHRVVMVPRYLNGSSDKNYAKALYTKAHI	201	
Db	142	ILVCECAPSKTGGLGVDVAGALPKALARGRHVVMVVPVLY-----GNVAPQHTGYRK	195	
Qy	202	KTPCGGSGHEVTFPFEYRDNDVWVEVDHPSY--HRPGSLYGNFGNAGFQNPQRYTLLCYAA	260	
Db	196	MFKIDQQDMENVNFIAYINDVFEDISPIFQHRGNNIYGGN---RVDLKRMDFUCKAA	252	

QY	261	CEAPLILELGGYIYG-NCMFVVVNDHWSLSLVPVLLAAKYPYGVYRDSRSTLVLIHNLAHQ	311
Db	253	IYVPWHVPGGGICYGCGNLVFIANDWHHTALLPVYLKAYFERDNGVMKFTRSVLVLIHNLIAHQ	312
QY	320	GVPEASTYPDGLUPPEWYGALFVPEWARRHALDKGEAVNELKGAVVVTADTRIVTSQGY	379
Db	313	GRGPMDDFISVDLUPAQYADLFKLYDPV-----GGDHFENIFAAGLKTADRVVTVSHCY	364
QY	380	SWEVTTAEGSGGQGNELLSSRKSVYNGVINGIDINDWNPTTKCLPH-----HYSVDDL-SG	434
Db	365	AWELKTSSEGGGLNGIRNENEMKLOQVINGIDITEENWPQLDYVLSKSDGYANYSLDTLQGT	424
QY	435	KAKCKAELQELPLVPREDVPLGFTGRDLYOKGIDLIKMAIPELMREDVQFVMLGSGDP	494
Db	425	KPOCKAALQKEMNLPRYDDVPLGFTGRDLHQKGVDLIAEALPMMQGDVQLVMLGTGRP	484
QY	495	IFEGNWRSTESSYKDKFRGVGVSFVSHRITAGCDILLMPSEPCGULNQLYAMQYGTIV	554
Db	485	DLEQMLKQIESQYGDVGRGVWVGSVKTARITAGADILLMPSEPCGULNQLYAMSYGTV	544
QY	555	PVHVGTLGGLRDVTETNPFCAKGEETGWAFSPLTYDKMLMALRTAMSTFRHKPSWESGL	614
Db	545	PVHAVGGLRDTYQPPDFP---NESGYGWTFGRAEANQLIDALGNCLLTIRYQKQSWESL	601
QY	615	MKRGMTKDHWTWDA 628	
Db	602	QRRGMQDLSDWDA 615	
RESULT 14			
Q48899	ID	PRELIMINARY; PRT; 732 AA.	
AC	O48899		
DT	01-JUN-1998 (T-EMBLrel. 06, Created)		
DT	01-JUN-1998 (T-EMBLrel. 06, Last sequence update)		
DT	01-JUN-2001 (T-EMBLrel. 17, Last annotation update)		
DE	STARCH SYNTHASE ISOFORM ZSTII-1 (BC 2.4.1.21) (STARCH (BACTERIAL		
DE	GLUCOGEN) SYNTHASE) (GLYCOGEN SYNTHASE) (ADP-GLUCOSE-STARCH		
DE	GLUCOSYLTRANSFERASE) (FRAGMENT).		
GN	SSLI OR ZSSIIA.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade		
OC	Panicoidae; Andropogoneae; Zea.		
OC	NCBI_TaxID=4577;		
RL	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W64A.; TISSUE=ENDOSPERM;		
RA	Knight M.E., Harrn C., Lilley C.E.R., Guan H., Singletary G.W.,		
RA	Mu-Forester C., Wasserman B.P., Keeling P.L.;		
RL	Plant J. 0:0-0(1998).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W64A.; TISSUE=ENDOSPERM;		
RA	Harrn C., Knight M., Ramakrishnan A., Guan H., Keeling P.L.,		
RA	Wasserman B.P.;		
RL	Plant Mol. Biol. 0:0-0(1998).		
CC	-1- CATALYTIC ACTIVITY: ADP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL) (N) = ADP-		
CC	+ ((1,4-ALPHA-D-GLUCOSYL)(N-1)).		
DR	EMBL; AF019296; AAD13341.1; -.		
DR	Mendel; 24658; Zeama; Sssl; 24658.		
DR	InterPro; IPR001296; Glycos_transf_1.		
DR	InterPro; IPR002086; Aldehyde_dehydr.		
DR	Pfam; PF00534; Glycos_transf_1.		
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.		
KW	Transferase; Glycosyltransferase.		
FT	NON_TER 732 732		
SEQ	SEQUENCE 732 AA; 80246 MW; 29A6745055090F3B CRC64;		

Query Match	30.4%;	Score 1227.5;	DB 10;	Length 732;
Best Local Similarity	40.3%;	Pred. NO. 1.2e-82;		
Matches 276:	Conservative 104;	Mismatches 212;	Indels 93;	Gaps 16;

OM of: US-09-674-824-2 to: EST:* out_format : pfs
 Date: Mar 28, 2002 4:55 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
 -Q/cgn2_1/USPTO.spool/US09674824/runat_27032002_094250_18407/app_query.fasta_1.824
 -DB=EST -OFMT=fastap -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000 -GAPOP=4.500
 -QAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi
 -LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
 -MINLEN=0 -MAXLEN=300000000 -USER=US09674824_CGNI_1_3637
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-674-824-2
 Query length: 756
 Database: EST:*
 Database sequences: 11351937
 Database length: 1077921985
 Search time (sec): 1499.200000

score.list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
gb_est2:BG343782	+ 1120.00	1305.81	1.4e-63	988	! Documentation
gb_est1:AV832913	- 944.00	1103.12	2.8e-52	706	! BG343782 HVSMEg0006L20f Hordeum
gb_est1:BE418354	+ 924.00	1076.68	8.3e-51	988	! AV832913 AV832913 K. Sato unpub
gb_est2:BF113316	+ 842.00	985.41	1.0e-45	536	! BE418354 SCL031.D08R990531 ITEC
gb_est1:BE434991	+ 807.00	944.62	1.9e-43	588	! BF113316 EST440906 tomato break
gb_est1:AW678087	+ 789.00	923.99	2.7e-42	562	! BE434991 EST406069 tomato break
gb_est1:AW932199	+ 735.50	861.41	8.1e-39	564	! AW678087 WSL_13.G10.bl.A002 Wat
gb_est1:AV832689	+ 710.00	833.57	2.9e-37	454	! AW932199 EST358042 tomato fruit
gb_est2:BF482660	+ 682.00	801.81	1.7e-35	408	! AV832689 AV832689 K. Sato unpub
gb_est1:AW759569	+ 643.50	753.03	8.8e-33	618	! BF482660 WHE2301-2304_E13_E132S
gb_est1:AW759946	+ 640.50	749.45	1.4e-32	623	! AW759569 si45a09.y1 Gm-cl027 G1
gb_est1:AW599325	+ 638.50	747.57	1.8e-32	592	! AW759946 si56b01.y1 Gm-cl027 G1
gb_est1:AU075475	+ 613.00	721.41	5.1e-31	396	! AW599325 gb13e02.y1 Moss EST li
gb_est1:AW438020	+ 555.00	653.27	3.2e-27	411	! AU075475 AU075475 Rice cDNA fr
gb_est2:BG050773	+ 543.00	639.47	1.9e-26	401	! AW438020 ST83F01 Pine Triplex S
gb_est2:BG607242	- 535.50	627.71	8.4e-26	557	! BG050773 FMI_70_P03.bl.A003 Fl
gb_est1:BE3453184	+ 517.50	605.98	1.4e-24	601	! BG607242 WHE2493.B06.D112S Tril
gb_est1:BE347900	+ 502.50	588.09	1.4e-23	625	! BE3453184 NF091H081F1073 Devel
gb_est1:BG350240	+ 496.00	578.03	4.3e-23	734	! BE347900 SP08603.y1 Gm-cl041 G1
gb_est1:BG529615	+ 495.50	580.07	3.8e-23	614	! BG350240 O86A01 Mature tuber la
gb_est2:BG351920	+ 494.00	576.75	5.8e-23	729	! AV524142 AV524142 Arabidopsis t
gb_est2:BG351175	+ 485.00	565.45	2.5e-22	794	! BG529615 EST504510 cSTS Solanu
gb_est1:BG351920	+ 484.00	565.45	2.5e-22	696	! BG351920 134D01 Mature tuber la
gb_est1:AV441053	+ 473.50	552.25	1.3e-21	773	! BE355965 sp96a10.y1 Gm-cl045 G1
gb_est1:AW569438	+ 468.00	548.29	2.3e-21	635	! BG351175 107B03 Mature tuber la
gb_est1:BF424294	+ 463.50	543.69	4.0e-21	589	! AV441053 AV441053 Arabidopsis t
gb_est1:BE414362	+ 460.00	537.44	9.0e-21	695	! AW569438 si87a06.y1 Gm-cl031 G1
gb_est2:BG456652	+ 459.00	535.53	1.0e-20	681	! BF424294 sr36f10.y1 Gm-cl051 G1
gb_est2:BG449742	+ 455.50	532.76	1.6e-20	754	! BE414362 SCU008.H03.R990714 Tr
gb_est2:BG350306	+ 448.50	530.63	2.2e-20	652	! BG456652 NF083D09PL1F1076 Phos
gb_est2:BI131080	+ 446.00	520.29	8.1e-20	335	! BG351232 088A10 Mature tuber la
gb_est1:AV18330	+ 443.50	517.28	1.2e-19	765	! CL9776 C19776 Rice panicle at r
gb_est2:BI406367	+ 439.00	520.71	7.7e-19	294	! BG350306 089D01 Mature tuber la
gb_est2:BG600852	+ 436.50	507.80	4.0e-19	882	! BI131080 EST5311930 GESP Medica
gb_est2:BI406798	+ 434.50	507.45	4.2e-19	709	! AV18330 AV18330 Lotus japonic
gb_est2:BI406798	+ 433.00	506.50	4.8e-19	649	! BI406367 167H04 Mature tuber la
gb_est2:BI406287	+ 432.50	505.50	5.4e-19	679	! BG600852 EST505747 cSTS Solanu
					! BI406798 180H08 Mature tuber la
					! BI406287 159B05 Mature tuber la

gb_est1:AW472190 + 426.00 499.73 1.1e-18 555 ! AW472190 si21b02.y1 Gm-cl029
 gb_est2:BF460244 + 419.00 488.78 4.6e-18 753 ! BF460244 073C04 Mature tuber
 gb_est1:AV638089 + 416.50 491.08 3.4e-18 424 ! AV638089 AV638089 Chlamydomo
 gb_est2:BG652154 + 415.50 488.02 5.1e-18 522 ! BG652154 sad74h10.y1 Gm-cl05
 gb_est2:BG350735 + 414.00 484.12 8.4e-18 661 ! BG350735 097D10 Mature tuber

seq_name: gb_est2:BG343782

seq_documentation_block:

LOCUS BG343782 988 bp mRNA EST 29-MAY-2001
 DEFINITION HVSMEg0006L20f Hordeum vulgare pre-anthesis spike EST library
 HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone
 HVSMEg0006L20f, mRNA sequence.
 ACCESSION BG343782
 VERSION BG343782.1 GI:13156111
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 988)

Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
 ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
 ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
 Wood,T.
 Development of a genetically and physically anchored EST resource
 for barley genomics
 Unpublished (2000)
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu

Seq primer: AATTACCCCTCACTAAGGG
 High quality sequence stop: 845.

FEATURES
 source
 1..988
 /organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMEg0006L20f"
 /clone_lib="Hordeum vulgare pre-anthesis spike EST library
 HVCNDA0008 (white to yellow anther)"
 /tissue_type="pre-anthesis spike"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
 more details on library preparation and sequence analysis
 see http://www.genome.clemson.edu/projects/barley/ To
 order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT 243 a 223 c 291 g 211 t 20 others

ORIGIN

alignment_scores:
 Quality: 1120.00 Length: 256
 Ratio: 4.786 Gaps: 5
 Percent Similarity: 91.406 Percent Identity: 88.281

alignment_block:

US-09-674-824-2 x BG343782 ..
 Align seg 1/1 to: BG343782 from: 1 to: 988

504 GluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValPr 520
 |||||
 3 GAGTCGAGCTACAAAGGATAAATTCGCGGATGGTGGATTAGTGTTC 52

520 oValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSera 537
 |||||
 53 AGTTTCCACAGAAATACTCGAGGTTCGATATATTGTTAATGCCATCCA 102

537 rgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThr 553
|||||
103 GATTGAAACCTTCGGCTCTTAATCAGCTATATGCAATATGGTACA 152
554 ValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPh 570
|||||
153 GTTCCTCTGTTCATGAACTGGGGCTTCGAGACACGGTGAGACCTT 202
570 eAsnProPheGlyAlaLysGlyGluGluGlyThrGlyThrAlaPheSerP 587
|||||
203 CAACCTTTTGGTCAAAAGAGAGAGGCTACAGGCTGGCGCTTCTCAC 252
587 roLeuThrValAspLysMetLeuThrAlaLeuArgThrAlaMetSerThr 603
|||||
253 CACTAACCGTGGAAAAAATGTTGGGCATTCGGAACCGCATTCGACG 302
604 PheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetTh 620
|||||
303 TTTAGGAGCACACAGCCCTCTGGAGGGCTCATGAAGCGAGGCATGAC 352
620 rLysAspHisThrTrpAspHisAla . ProSerSerThrSerArgSerSer 636
|||||
353 GAAAGACCATACGTGGGACCATGCGCGCGAGCAGTACGAGCAGATCTCG 402
637 SerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuGlyAr 653
|||||
403 AGTGGGCTTCGTGGACCAACCTACGTATGATAGCTGGGACTGGGAG 452
653 qSerLysCysGluSerProSerAlaLeuLysThrSerSerSerPheA 670
|||||
453 GTCGAAGCGCCAGCTCTCTGAGCTCTCAAGACATGTTCTCTCATCTCC 502
670 rgGlyProGlu... GlyTyrProCysThrLeuArgCysProAlaThrVal 685
|||||
503 CGGGCCGGAAGGAGATACCCCTGTACATTGCGTTGCTGCTACAGTA 552
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702
|||||
553 GAGTCGCAATGGCTGTGCTGCT . TGGTTCGCCGGTTCGAGAGTAGACGA 601
702 pGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGln . LeuGln 718
|||||
602 CGGTGCAACGCTGCGCG... GCAGCTTCGGGTGGATGACAGNNNTACAG 648
719 PheTrpGlyLeuArgLysCysAlaAlaGlyTrpLeuThrAlaLysHi 735
|||||
649 TTTTGGGGAATAAGGAAGGAGTCTCGGTGGGTGTACAGCTGAAATCAGAAAC 698
735 shiSerAspGlySer . LeuSerValArgValThrAlaGluLeuArgAsn 751
|||||
699 CACCTCAGTCGCGAGCTCTCTCGGTGGGTGTACAGCTGAAATCAGAAAC 748
752 GlnLeuValThr 755
|||||
749 CCACCTGGTGTCT 760

seq_name: gb_est1:AV832913

seq_documentation_block: 706 bp mRNA EST 22-JUN-2001
LOCUS AV832913
DEFINITION AV832913 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare top three leaves adult, heading stage Hordeum vulgare subsp. vulgare cDNA clone rbaal2b23, mRNA sequence.

ACCESSION AV832913
VERSION AV832913.1 GI:14525002
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 706)

AUTHORS
TITLE
JOURNAL
COMMENT

Sato, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2001)
Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kassato@rib.okayama-u.ac.jp,
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct
submission:
database: http://www.shigen.nig.ac.jp/barley/Barley.html.
Location/Qualifiers

FEATURES

source
1..706
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="rbaal2b23"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare top three leaves adult, heading
stage"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
BASE COUNT 174 a 213 c 162 g 155 t 2 others
ORIGIN

alignment_scores:

Quality: 944.00 Length: 214
Ratio: 4.744 Gaps: 4
Percent Similarity: 92.991 Percent Identity: 90.187

alignment_block:

US-09-674-824-2 x AV832913/rev ..
Align seg 1/1 to reverse of: AV832913 from: 1 to: 706

546 LeuTyrAlaMetGlnTyrGlyThrValProValHisGlyThrGlyGlu 562
|||||
705 CTATATGCTATGCAATATGGTATGATGCTCTCTGTTGTTTCATGGAACCTGGGG 656
562 YLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyGluG 579
|||||
655 CTTTCGAGACACGGTGAGACCTTCAACCTTTTGGTGCAGAGGAGGAGG 606
579 luGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLeuTrp 595
|||||
605 AGGCTACAGGGTGGCGTCTCACCACCTAAACCGTGGAAAAAATGTTGTGG 556
596 AlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGlu 612
|||||
555 GCATTGGCAACCGCATTTTCAGCAGTTTAGGGAGCAGACAGCCCTCTCTGGGA 506
612 uGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla . 628
|||||
505 GGGGCTCATGAAGCGAGCATGACGAAAGACCATACGTGGGACCATGCGC 456
629 ProSerSerThrSerArgSerSerSerGlyProSerTrpThrAsnProTh 645
|||||
455 CCGAGCATACGAGCAGATCTTCGAGTGGGCTTCGTGGGACCAACCTTAC 406
645 rSerCysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaAl 662
|||||
405 GTCATGTAGACTGGGACTGGGAGGTCCAGGCGCAGTCTCTTTGAGCTC 356
662 euLysThrSerSerSerPheArgGlyProGlu... GlyTyrProCys 677
|||||
355 TGAAGACATGTTCTCTCATCTTCGCGCGCGGAGAGGAGATACCCCTGT 306
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysAlaLeuLeuTr 694
|||||
305 ACATTGCGGTGTGCTGTACAGTAGAGTGCAGTCCGCTTGTGCTTGT.TG 257

694 pPhealaglySerArgThrTrpAspGlyCysAlaAlaAlaValThrA 711
 256 GTTCGCGGTTCCGAGTAGACGAGGTCGCAACGCTGCCG...GCAG 210
 711 laSerGlyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727
 209 CTTCCGGGTGGATGACAGTAGTATTTGGGGAATAAGGAGGAGGTGCT 160
 728 AlaGlyTrpLeuThrAlaLysHisHisSerAspGlySer..LeuSerVal 743
 159 GCAGGATGTTGACAGCAACACACCTCAGTCGCGACGCTCTCTCGGTG 110
 744 ArgValThrAlaGluIleArgAsnGlnLeuValThrLeu 756
 109 GGTGTTACAGCTGAATCAGAAACCACTGGTGCTCTT 71

seq_name: gb_est1:BE418354

seq_documentation_block: 988 bp mRNA EST 24-JUL-2000
 LOCUS BE418354
 DEFINITION SCL031.D08R90531 ITEC SCL Wheat Leaf Library Triticum aestivum
 cDNA clone SCL031.D08, mRNA sequence.
 ACCESSION BE418354
 VERSION BE418354.1 GI:9416200
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Triticum.
 1 (bases 1 to 988)

REFERENCE
 AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier
 S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
 Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
 Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
 Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M.,
 Sorrells, M., Warburton, M., and Wenzel, G.
 International Triticeae EST Cooperative (ITEC): Production of
 Expressed Sequence Tags for Species of the Triticeae
 Unpublished (2000)
 CONTACT: Cloutier S
 Cereal Research Centre, Agriculture & Agri-Foods Canada
 Winnipeg, MB, CANADA
 Tel: 204 983 2340
 Fax: 204 983 4604
 Email: scloutier@em.agr.ca
 International Triticeae EST Cooperative (ITEC)
 http://wheat.pw.usda.gov/genome.

FEATURES
 source
 1..988
 /organism="Triticum aestivum"
 /cultivar="Thatcher Lrl"
 /db_xref="taxon:4565"
 /clone="SCL031.D08"
 /clone_lib="ITEC SCL Wheat Leaf Library"
 /tissue_type="young leaf"
 /dev_stage="after 24 hour challenge with LR-AVR1"
 /note="vector: Lambda ZAP; 1.0 kbp average insert size."
 227 a 199 c 244 g 247 t 71 others

BASE COUNT
 ORIGIN
 227 a 199 c 244 g 247 t 71 others

alignment_scores:
 Quality: 924.00 Length: 225
 Ratio: 4.738 Gaps: 9
 Percent Similarity: 86.667 Percent Identity: 84.889

alignment_block:
 US-09-674-824-2 x BE418354 ..

Align seg 1/1 to: BE418354 from: 1 to: 988

475 AlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGlyse 491

110 GCCATTCAGAGCTCATGAGGAGGAGCGTCAATTTGTCTCATGCTTGATC 59
 491 rGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrL 508
 60 TGGGATCCCAATTTTGAAGGCTGGATGATATCTACCGAGTCGAGTTACA 109
 508 ysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArg 524
 110 AGGATAAATTCCTGGATGGTGGATTAGTGTCTCCAGTTTCCACAGA 159
 525 IleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCy 541
 160 ATAACGTCAGGTGGCGATATATTGTTAATGCCATCGAGATTTTGAACCTG 209
 541 sGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThr.ValProValVal 557
 210 CGGTCTTAATCAGCTATATGCTATGCAATATGTCACACGTTCTCTGATGTT 259
 558 HisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574
 260 CATGGAACCTGGGGCCCTCCGAGACACAGTCGAGACCTTCAACCTTTGG 309
 574 yAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValA 591
 310 TGC AAAAGAGAGAGAGGAGGTACAGGGTGGGGTCTCACCGCTAACCGTGG 359
 591 sPlysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHis 607
 360 ACAAGATGTTGTGGGCATTTCCGAACCGCATGTGCGACATTCAGGAGAC 409
 608 LysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisTh 624
 410 AAGCCCTTCTGGAGGGGCTNATNAAGCAGGAGCATGACGAAAGACCATAC 459
 624 r.TrpAspHisAla..ProSerSerThrSerArgSerSerSerGlyProse 640
 460 CGTGGGACCATCGCCCGAGCAGTAGCAGACAGATNTTTGAGTGGGCTTC 509
 640 rTrpThrAsnProThrSer.CysArgArgGlyLeu.GlyArgSerLysCy 656
 510 GTGGACCAACCTTACNTCATTTAGACNGGAGCTGGGAGAGTCCAAAGTG 559
 656 sGluSerProSer..AlaLeuLysThrSer.SerSerSerPhe....ArgGl 671
 560 CGAGTCTCTTTGAGCTTTGAANACATCCCTCTTTATTTCTTTNCGCGGNC 609
 671 yProGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerG 688
 610 CCGGGAAGGATACCCCTGTNCATT..... 634
 688 lncysalacys 691
 635 ..TGCTTTGT 643

seq_name: gb_est2:BF113316

seq_documentation_block:

LOCUS BF113316 mRNA EST 18-MAY-2001
 DEFINITION EST440906 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLEGA4E24 5' sequence, mRNA sequence.

ACCESSION BF113316
 VERSION BF113316
 KEYWORDS EST.
 SOURCE tomato.

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE 1 (bases 1 to 596)

AUTHORS Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,

Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley
S.D.

Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)

CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES

source

1..596
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_lib="tomato breaker fruit"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCuadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

BASE COUNT 167 a 120 c 149 g 160 t
ORIGIN

alignment_scores:
Quality: 842.00 Length: 198
Ratio: 4.626 Gaps: 0
Percent Similarity: 91.919 Percent Identity: 75.758

alignment_block:
US-09-674-824-2 x BF113316 ..
Align seg 1/1 to: BF113316 from: 1 to: 596

293 LeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgSerArgSe 309
|||||
1 CTTTATGAGCAGCAGTATCGTCTTATGCTTTACAGGATGCTCGTAG 50
309 rThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProLase 326
| : : : : :
51 TATTGTCGCAATACACAACTTGCATCAGGAGTGGAGCTCGACGAA 100
326 hrTyrProAspLeuGlyLeuProProGluTyrGlyAlaLeuGluTrp 342
||||| : : : : :
101 CCTACAATAATTGGGATGGCTCTCTCACTGATGGAGCACTTGAATGG 150
343 ValPheProGluTrpAlaArgArgHisAlaLeuAspLysGlyAlaAla 359
: : : : :
151 ATATTTCACATGGCAGGAGGCCCTGCTGACACTGTGTGAACAGT 200
359 IasnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrVal 376
||||| : : : : :
201 GAATGTTTGAAGGGCAATCTCAGTTGCTCATCGGATCTGACAGTTA 250
376 erGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGlnGlyLeu 392
||||| : : : : :
251 GCCAGGGATACATCGGGAATAACAACTCTGGAAGGGGATATGGGCTA 300
393 AsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnG 409
: : : : :
301 CATGAGCTGCTGAGCAGTAGACAGTCAGTTCTTAATGGAATTAATAG 350
409 yIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuProHis 426
||||| : : : : :
351 AATAGATGTTAATGATGGACCCCTCGACAGATGAGCATATGCTTCGC 400
426 iSyrSerValAspLeuSerGlyLysAlaLysCysLysAlaGluLeu 442
: : : : :
401 ATTACTCCATCAATGACCTCTCTGAAAGGCTCAGTGCAGAGCATCTG 450

443 GlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPh 459
|||||
451 CAAAGGAAGTGGCCTTCCAAATTCAGCTGATTCGCTGATTCGCTGAT 500
459 eIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAla 476
|||||
501 TATTGGAAGCTGGAGTACCAGAAAGGTGTTCACATAATATCTGTCAG 550
476 leProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490
|||||
551 TTCCAGAACTTTTCAGAGAGGATGTCCTCAATTTGTAATGCTTGA 594

seq_name: gb_est1:BE434991

seq_documentation_block: 588 bp mRNA EST 18-MAY-2001
LOCUS BE434991
DEFINITION EST406069 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone CLEG24J21, mRNA sequence.

ACCESSION BE434991
VERSION BE434991.1 GI:9432834
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 588)
AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley,
S.D.

TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: CUGI
Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

source 1..588
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_lib="CLEG24J21"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCuadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

BASE COUNT 174 a 115 c 149 g 150 t
ORIGIN

alignment_scores:
Quality: 807.00 Length: 195
Ratio: 4.585 Gaps: 0
Percent Similarity: 90.256 Percent Identity: 73.333

alignment_block:
US-09-674-824-2 x BE434991 ..

Align seg 1/1 to: BE434991 from: 1 to: 588

311 LeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThry 327
: : : : :
3 GTCCGAATACACAACATTTGCACATCAGGGAGTGGAGCTGCAGCAACCTA 52

509 TGCCGCTGAACAATACGAACAATCTTCCAGTGGGCTTCATCGATCGAC 558

seq_documentation_block:					
LOCUS	AW678087	562 bp	mrna	EST	19-JUL-2000
DEFINITION	WS1_13_G10_b1_A002 Water-stressed 1 (WS1) Sorghum bicolor CDNA, mRNA sequence.				
ACCESSION	AW678087				
VERSION	AW678087.1	GI:7551807			
KEYWORDS	EST.				
SOURCE	sorghum.				
ORGANISM	Sorghum bicolor				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.				
AUTHORS	1 (bases 1 to 562) , L.H. Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt				
TITLE	An EST database from Sorghum: water-stressed plants				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Cordonnier-Pratt MM Department of Botany The University of Georgia Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 542 1805 Email: mmpratt@uga.edu Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence				

/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
BASE COUNT 112 a 84 c 126 g 130 t 2 others
ORIGIN

alignment_scores:
Quality: 710.00 Length: 143
Ratio: 5.035 Gaps: 0
Percent Similarity: 98.601 Percent Identity: 95.105

alignment_block:
US-09-674-824-2 x AV832689 ..

Align seg 1/1 to: AV832689 from: 1 to: 454

106 GlyIleAlaGluAspSerIleAspSerIleValAlaAlaSerGluGI 122
|||||
27 GGGATTGCTGAGGATTCCTC.GACACCATAGTTGTGGCTGCAAGTGAGCA 75
122 nAspSerGluIleMetAspAlaAsnGluInProGlnAlaLysValThrA 139
76 GGATTCTGAGATCATGGATGCCAACCAACCTCTAGCTAAAGTTACAC 125
139 rgSerIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGly 155
126 GTAGCATGGTGTGTGACTGGTGAAGCTGCTCTTATGCAAGTCAGGG 175
156 GlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaArgGI 172
176 GGGCTGGGAGATGTTGTGGTCTGCCAATGCTCTTGTCTGCTCGTG 225
172 yHisArgValMetValValMetProArgTyrTyrLeuAsnGlySerSerAspL 189
226 TCACCGTGTGATGTTGTTAATGCCGAGATACTTAAATGGGACCTCTGATA 275
189 yAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCys 205
276 AAAACTATGCAAGGCATTATACACCGGAAGCACATTAAGATTCCATGC 325
206 PheGlyGlySerHisGluValThrPheHisGluTyrArgAspAsnVa 222
326 TTCGGGGATCACATGAAGTCACCTTTTTCATGAGTATAGACACACGT 375
222 LaspTrpValPheValAspHisProSerTyrHisArgProGlySerLeut 239
376 CGATTGGGTGTTTGTGATCATCATCATATCATATAGACACCAAGTTGT 425
239 yrGlyAspAsnPheGlyAlaPheGlyAsp 248
426 ATGGAGATAATTTTGGTCTTTTGGTGAT 454

seq_name: gb_est2:BF482660

seq_documentation_block:
LOCUS BF482660 408 bp mRNA EST 06-DEC-2000
DEFINITION WHE2301-2304_E13_E13Zs wheat pre-anthesis spike cDNA library
Triticum aestivum cDNA clone WHE2301-2304_E13_E13, mRNA sequence.
ACCESSION BF482660
VERSION BF482660.1 GI:11565884

KEYWORDS
SOURCE EST.
bread wheat.
Triticum aestivum

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 408)

REFERENCE
AUTHORS
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)

COMMENT

Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stragene SK primer.

FEATURES

Source

1..408
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE2301-2304_E13_E13"
/clone_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give phluescript
phagemids in the TJ Close lab (Choi, Close, Fenton) at
the University of California, Riverside. plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 108 a 76 c 99 g 125 t
ORIGIN

alignment_scores:

Quality: 682.00 Length: 136
Ratio: 5.128 Gaps: 0
Percent Similarity: 97.794 Percent Identity: 94.118

alignment_block:

US-09-674-824-2 x BF482660 ..

Align seg 1/1 to: BF482660 from: 1 to: 408

427 TyrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGI 443
1 TATTCGTGCGACCTCTCTGGAAGGCCAAATGTAAGCTGAATTGCA 50
443 nLysGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheI 460
51 CAAGGAGCTGGCTTTACCTGTAAAGGAGGATGTTCTCTGATTGGCTTTA 100
460 leGlyArgLeuAspTyrGlnLysGlyIleAspLeuLysMetAlaIle 476
101 TTGGAGAGCTGGATTACCAAGGCATTGATCTCATTAATAATGCCATT 150
477 ProGluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAs 493
151 CCAGAACTACTGAGGAGGACGTACACTTTGTCATGCTGGATCTGGGGA 200
493 pProfilePheGluGlyTrpMetArgSerThrGluSerTyrLysAspL 510
201 TCCAAATTTTGAAGGCTGGATCGAGATCTACCGAGCTAGTTACAGGATA 250
510 ysPheArgGlyTrpValGlyPheSerValProValSerHisArgIleThr 526
251 AATTCGCGGATGGGTGGATTACTGTCCTCCCAATTTCCACACAATAACT 300
527 AlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLe 543
301 GCATGTTCCGATATATGTTAATGCCATCCATGTTTGAACCTTTGGGTCT 350
543 uAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValHisGlyT 560

|||||

351

TAATCAACTATATGCTATGCAATATGTTACAGTTCCTGTAGTCATGGAA

400

560

hrGlyGly 562

|||||

401

CTGGGGGC 408

seq_name: gb_est1:AW759569

seq_documentation_block: 618 bp mRNA EST 18-JUL-2000

LOCUS AW759569

DEFINITION sl45a09.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl027-4241 5' similar to SW-UGS3_PEA Q43093 GLYCOGEN [STARCH] SYNTHASE PRECURSOR ; , mRNA sequence.

ACCESSION AW759569

VERSION AW759569.1 GI:7691442

KEYWORDS soybean.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 618)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,I., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Insert Length: 1230 Std Error: 0.00
High quality sequence stop: 411.

FEATURES

source

Location/Qualifiers

1..618

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-4241"

/clone_lib="Gm-cl027"

/tissue_type="cotyledons of 3- and 7-day-old Williams seedlings"

/lab_host="DH10B"

/note="vector: pbluescript II SK+; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 3- and 7-day-old Williams seedlings which were propagated on paper towels with distilled water. The cotyledons were flash-frozen in liquid nitrogen, then lyophilized for 72 hours. Unequal amounts of mRNA was used for cDNA synthesis. StrataGene's cDNA Synthesis kit (catalog number 200401) was used to synthesize the cDNA. First- stranded synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An anchor nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGACTAGTCTCGAG(T)18] to anchor the primer at the 5' end of the poly(A) tract. After second- strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the

first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into StrataGene's pBluescript(tm) II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 179 a 115 c 168 g 152 t 4 others

ORIGIN

alignment_scores:

Quality: 643.50 Length: 203

Ratio: 3.900 Gaps: 2

Percent Similarity: 81.281 Percent Identity: 63.054

alignment_block:

US-09-674-824-2 x AW759569 ..

Align seg 1/1 to: AW759569 from: 1 to: 618

369 AlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrTh 385

1 GCTGACCCGGATTGTCACCTGTGAGTCATGGATATGCATGGGAGATTAAAC 50

385 rAlaGluGlyClnGlnGluLeuLeuLeuSerSerArgLysSerV 402

51 TTCTGGAAGTGGTGGGGTTGGCATGGGATCAATAATGAGAATGACTGGA 100

402 alLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 418

101 AATTGACAGGAGTGTGCAATGGATTTCACACCAAGATGGAACCCCAAG 150

419 ThrAspLysCysLeuProHis.....HisTyrSerValAspAs 431

151 ATTGATGTTCACTTGAATCAGATGATGATCACTAACTACACCTTGAGAC 200

431 pLeu...SerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuG 447

201 CCTCAAAAGTGGCAAGCGTCAGTCACAAAGCCGCTTCGAAAGGAGCTCG 250

447 lyLeuProValArgGluAspValProLeuIleGlyPheIleGlyArgLeu 463

251 GTTTGCCTGTCCTGAGGATGTTCCGTACTTGGATTTCATTGGAAGGCTG 300

464 AspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProGluLeuMe 480

301 GATCAACAGAAGGCAATGATCTCATGCCCAAGCAATTCCTTGGATAGT 350

480 tArgGluAspValGlnPheValMetLeuGlySerGlyAspProIlePheG 497

351 GGGCCAGATGTGCAGCTAGTCATGTTGGAACTGGGAAGCGCGGACTTAG 400

497 luGlyTrpMetArgSerThrGluSerSerTyrLysAspLysPheArgGly 513

401 AAGATATGCTTANGCAGTTTGTGATCCCAACCCGTCGACAAAAGTCAGAG 450

514 TrpValGlyPheSerValProValSerHisArgIleThrAlaGlyCysAs 530

451 TGGGTGGCTTTTCNGTCAAGATGGCTCACCNGATAACAGCANGTGCAGAA 500

530 pileLeuLeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeuT 547

501 CATATTGCTGATGCCATCAAGATTGTCGACCATGTGATTGAATCAACTCT 550

547 yAlaMetGlnTyrGlyThrValProValValHisGlyThrClyGlyLeu 563

564 ArgAspThr 566
 ::::::::::
 606 AAGGATACA 614

seq_name: gb_est1:AW599325

seq_documentation_block: 592 bp mRNA 22-MAR-2000
 LOCUS AW599325 EST
 DEFINITION gb13e02.y1 Moss EST library PPN Physcomitrella patens cDNA clone
 PEP_SOURCE_ID:PPN091503 5' similar to TR:O24398 O24398 STARCH
 SYNTHASE ;, mRNA sequence.

ACCESSION AW599325
 VERSION AW599325.1 GI:7286838
 EST
 SOURCE Physcomitrella patens.
 ORGANISM Physcomitrella patens.

REFERENCE 1 (bases 1 to 592)
 AUTHORS Quatrano, R., Bashlades, S., Cove, D., Cuming, A., Knight, C., Clifton
 S., Marra, M., Hillier, L., Pape, D., Martin, J., Wyllie, T., Underwood
 K., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T.,
 Steptoe, M., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
 Waterston, R. and Wilson, R.
 Leeds/Wash U Moss EST Project
 Unpublished (1999)
 TITLE Contact: Ralph Quatrano
 JOURNAL Leeds/Wash U Moss EST Project
 COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Libraries were constructed by Dr. Stavros Bashlades as part of the
 Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
 Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
 University Genome Sequencing Center For information on obtaining a
 clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 419.

FEATURES
 source
 1..592
 /organism="Physcomitrella patens"
 /db_xref="taxon:3218"
 /clone="PEP_SOURCE_ID:PPN091503"
 /clone_lib="Moss EST library PPN"
 /tissue_type="protonemata: 7 day old tissue auxin treated"
 /lab_host="Dhl08"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; Construction of the cDNA library was carried out
 using Stratagene's 'Unizap - cDNA synthesis kit'. cDNA was
 constructed using an oligo dt primer/linker that contains
 a XhoI site within it. Following ds cDNA synthesis,
 EcoRI adapters were ligated to the blunt ends and sample
 was digested with XhoI. The result is cDNA with an EcoRI
 sticky end on one side and a XhoI sticky end on the other.
 This cDNA was ligated directionally in Unizap arms. The
 vector is designed containing the pBluescript sequence as
 well as lambda DNA and cDNA is cloned within this
 pBluescript sequence. The vector was then packaged using
 Gold gigapackaging extracts. Library was grown in XLBlue
 MRF' cells and amplified. The library was excised by mass
 excision using Stratagene's 'Mass excision kit' that uses
 exsist as a helper phage that releases the pBluescript
 sequence and circularises it as single stranded plasmids
 that are then packaged (by helper phage) and secreted out
 of the host cell as phagemids. SOLR cells were transformed
 with phagemids and the library was plated out on LB-amp
 plates to select for transformants. Approximately 1,000
 ,000 colonies were grown and recovered. The double
 stranded plasmid library was recovered by using Qiagen
 Midi prep kit. 2 micro grams of each library were used to

BASE COUNT 142 a 126 c 161 g 160 t 3 others
 ORIGIN

alignment_scores:
 Quality: 638.50 Length: 199
 Ratio: 3.893 Gaps: 2
 Percent Similarity: 82.412 Percent Identity: 61.307

alignment_block:
 US-09-674-824-2 x AW599325 ..

Align seg 1/1 to: AW599325 from: 1 to: 592

427 TyrSerValAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGl 443
 ::::::::::
 11 CACTCTTTAGAC.....ACAGGTAAGGCGCATGCAAGGCTGCTTTACA 54
 443 nLysGluLeuGlyLeuProValargGluAspValProLeuIleGlyPheI 460
 ::::::::::
 55 AAGGGAGTTGGGCTTGGCGGTGGCGGTGACGTCCCATTCCTTGGATTCA 104
 ::::::::::
 460 leGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIle 476
 ::::::::::
 105 TAGGTCCCTTGTATCACCACCAAGGTTATTGACATAATAGGCCAAGCTATG 154
 ::::::::::
 477 ProGluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAs 493
 ::::::::::
 155 CCATGGATGATGGACCAAGATATCCAACTGGTAATGCTCGGCACTGGCCG 204
 ::::::::::
 493 pProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspL 510
 ::::::::::
 205 GAAAGATTATTAAGACATGTTAGGCGCATTTGAAGGTTCTCATCGAGACA 254
 ::::::::::
 510 ysPheArgGlyTrpValGlyPheSerValProValSerHisArgIleThr 526
 ::::::::::
 255 AAGTTCGGGGATGGTTGGTTTCTGTGACCACTTCCCATCGGATTACC 304
 ::::::::::
 527 AlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLe 543
 ::::::::::
 305 GCAGGTGTTGATATTTGTTGATGCCATCTCGGTTTGAACCCCTGTGGCTT 354
 ::::::::::
 543 uAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyT 560
 ::::::::::
 355 GAACCAACTGTATGCGATGCGGTATGCGACAGTCCAGTAGTGCATCGTG 404
 ::::::::::
 560 hrGlyGlyLeuArgAspThrValGluThrPheAsnPropheGlyAlaLys 576
 ::::::::::
 405 TTGTGGGCTGAAAGATACAGTGCAGTCGTCGTTCAATCCCTTC..... 445
 577 GlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMe 593
 ::::::::::
 446 AACGAGTCAGGCGCTANGATGGAACNTTCGAAACACTACATGTGGACGCTT 495
 593 tLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProS 610
 ::::::::::
 496 TATCATCCCCCTTGGGAATGCAATTTGGACCTACCGTGCATTCAAGACTA 545
 610 erTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrp 625
 546 GTTGAAGGCGCATCAGCAACGATGATGTCTCTCANGACTTGAGCTGG 592

seq_name: gb_est1:AU075475

seq_documentation_block:
 LOCUS AU075475 396 bp mRNA EST
 DEFINITION AU075475 Rice cDNA from immature leaf including apical meristem
 Oryza sativa cDNA clone E60759_1A, mRNA sequence.
 ACCESSION AU075475
 VERSION AU075475.1 GI:5056096
 KEYWORDS .EST.
 SOURCE Oryza sativa.

ORGANISM

Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 396)
Sasaki, T. and Yamamoto, K.
Rice cDNA from immature leaf including apical meristem
Unpublished (1997)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@agr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP',

FEATURES

source
1..396
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="P60759.1A"
/clone_lib="Rice cDNA from immature leaf including apical
meristem"
/dev_stage="immature"
/note="Organ: leaf; Immature leaf including apical
meristem (under long day condition)."
BASE COUNT 104 a 72 c 103 g 116 t 1 others
ORIGIN

alignment_scores:

Quality: 613.00 Length: 132
Ratio: 4.827 Gaps: 1
Percent Similarity: 96.212 Percent Identity: 88.636

alignment_block:

US-09-674-824-2 x AU075475

Align seg 1/1 to: AU075475 from: 1 to: 396

440 AlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLe 456
2 GCTGAATTCGAGAGGAGCTGGGTTTACCTATAAGGCCCGATGCTCT 51
456 uileGlyPheilleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleL 473
52 GATTGGCTTTATTGGAAGATTGGACTATCAAAAGGCATTGATCTAATTA 101
473 yMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeu 489
102 AACTTGCATTCAGATCTCATCGGAGCAATATTCATTCGTATGCTT 151
490 GlySerGlyAspProIlePheGluGlyTyrMetArgSerThrGluSerSe 506
152 GGATCTGGTGACCCAGGTTTGAAGATGATGATGATGATGATGATGATG 506
506 rTyrLysAspPheArgGlyTrpValGlyPheSerValProValSerH 523
202 GTACAGGGATAAATTCGTGGATGGGTGGATTTAGTTCCTCCAGTTTCC 251
523 isArgIleThrAlaGlyCysAspIleLeuMetProSerArgPheGlu 539
252 ACCGATACTGCAGGTTCGATATATTTGATGATGATGATGATGATGATG 301
540 ProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVa 556
302 CCTTGTGGCTCAATCAGCTATATGCTATGCAATATGCTATGCTATGCT 351
556 lValHisGlyThr..GlyGlyLeuArgAspThrValGluThrPhe 570
352 TGTTCATGGGAAGCTGAGAGGCTCTCAGATACAGTACAGTGGAGATTTT 396

seq_name: gb_est1:AW438020

seq_documentation_block:

LOCUS AW438020 411 bp mRNA EST 14-FEB-2000
DEFINITION ST83F01 Pine triplex shoot tip library Pinus taeda cDNA clone
ST83F01, mRNA sequence.

ACCESSION AW438020
VERSION AW438020.1 GI:6973326
KEYWORDS EST.
SOURCE loblolly pine.

ORGANISM Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 411)

REFERENCE Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten

Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801

Email: rosswhetten@unity.ncsu.edu
Seq primer: 5' lambda triplex2 Sequencing Primer.

FEATURES

Location/Qualifiers
1..411
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST83F01"
/clone_lib="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site 1:
Sfil (A); Site 2: Sfil (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."

BASE COUNT 111 a 78 c 95 g 113 t 14 others
ORIGIN

alignment_scores:

Quality: 555.00 Length: 137
Ratio: 4.549 Gaps: 0
Percent Similarity: 89.051 Percent Identity: 72.993

alignment_block:

US-09-674-824-2 x AW438020

Align seg 1/1 to: AW438020 from: 1 to: 411

405 GlyIleValAsnGlyIleAspIleAsnAspTyrAspProThrThrAspLy 421
1 GCATCACCAGATGGATTGATGTTGATGATGATGATGATGATGATGATG 50
421 sCysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAlaLysC 438
51 ACATGTCCCATTCCTACTCTGTAGAGGACCTTCTGGAAGGCAAAAT 100
438 yslsAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspVal 454
101 GCAGGACGCTCTACAATGGAGCTGGACTTCTGTTCCAGCATGTTG 150
455 ProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLe 471
151 CGCGTATGTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 200
471 uileLysMetAlaIleProGluLeuMetArgGluAspValGlnPheValM 488

THIS PAGE BLANK (USPTO)